UNIVERSAL ELECTROMAGNETICS LTD

Variants of Gustavsen and Semlyen's Vector Fitting Algorithm for Impedance Function Characterisation

UELtd/AJM/PV/Zed_estimation/v1.0

Cover + viii + 75 pages

November 2021

UNIVERSAL ELECTROMAGNETICS LTD
6 THE RUGG
LEOMINSTER
HEREFORDSHIRE HR6 8TE, UK

EMAIL: andrew@ueltd.co.uk
WEBSITE: www.ueltd.co.uk

Although the author has made every reasonable effort to ensure the accuracy of the contents of this document, neither the author nor Universal Electromagnetics Ltd. will be held liable for damages resulting from the use or application of the material contained, to the extent permitted by international law.

This document may be distributed freely subject to the following conditions:

The document may not be changed or modified or the contents altered. It must remain complete without selective removal of material. It may not be sold or reproduced for commercial gain, without the written approval of an authorised representative of Universal Electromagnetics Ltd. (UK).

Universal Electromagnetics Ltd.
6 The Rugg, Leominster,
Herefordshire, HR6 8TE, UK.

Author Dr Andrew Mackay

Date November 2021

Issued by Universal Electromagnetics Ltd.
6 The Rugg
Leominster
Herefordshire HR6 8TE
UK.

Document changes record

Issue	Date	Change summary
v1.0	November 2021	First release.

Executive Summary

This report is concerned with variants of Gustavsen's vector fitting algorithm for the location of poles and residues of a sampled impedance function specified at real frequencies. Both Gustavsen and Semlyen's original variant and Gustavsen's second variant are considered in detail, and further modifications are also described. One of these modifications, referred to as the null-space method, shows particular promise and complements the first variant in its regions of applicability. It can be used in situations where the model order number is much larger than would otherwise be possible. This is of particular use in the finite order modelling of transmission line problems of infinite order.

A number of examples are selected, both from Gustavsen's original papers and in the characterisation of a radar absorbent material. The different variants are compared and contrasted when the order number is both over and underestimated and in the presence of significant noise. Two error measures are employed, E_1 and E_2, the former employed in Gustavsen and Semlyen's original method and the latter defining the residual error between the model and the original data. It is shown that the latter, E_2, provides a more reliable measure of accuracy. Although E_1 and E_2 can be equal this is not always true, even when the methods converge.

Conventions

In this report j will be used to represent the square root of -1 in keeping with common electrical engineering practice. All analysis at fixed frequency ω will assume a harmonic time convention with time dependence, $\exp(j\omega t)$. The Laplace transform variable $s = j\omega$.

List of contents

Document changes record iv

Executive Summary v

Conventions vi

List of contents vii

1 Project information and administration 1

2 Formulation 2
2.1 Requirements 2
2.2 Frequency and impedance scaling and the reflection coefficient 3
2.3 An outline of Gustavsen's first and second variant 3
2.4 Summary of new variants of the method 6
2.5 The single null-eigenvalue method 7
2.6 The null-space method 11
2.7 Gustavsen's first and a modification of Gustavsen's second variant 12
2.8 Gustavsen's second variant 14
2.9 A note on the error measures in the different methods 18

3 Numerical Examples 20
3.1 Summary 20
3.2 Example 1. A simple 5-pole response in the absence of noise 21
3.3 Example 2. Single real pole in the presence of noise 23
3.4 Example 3a. Gustavsen 18 pole example 25
3.5 Example 3b. Gustavsen 18 pole example with PR residues 27
3.6 Example 4. Gustavsen's 18 pole example with model order overestimated 29
3.7 Example 5. The 18 pole example with model order underestimation 37
3.8 Example 6. The 18 pole example with model order overestimation and noise 47
3.9 Interim conclusions and application to radar absorbent materials 55
3.10 Conclusions 68

4 References 69

Appendices 70

A **Least squares minimisation of the residual error E_2** **70**

B **Alternative proof of Gustavsen's matrix method for the zeros of $\sigma(s)$** **73**

1 Project information and administration

The work in this report was conducted by the author for Universal Electromagnetics Ltd. It is private venture work and is the sole property of the author. No third party contains the rights to any intellectual property contained within.

2 Formulation

2.1 Requirements

It is frequently the case that an impedance function, $Z(\omega)$, assumed to be passive and linear, can be evaluated over an arbitrary range of real frequencies, ω, but analysis requires a representation as a function of the Laplace variable $s = j\omega$. For example, an unknown network may present a measurable driving point impedance or analysis software may provide a complex reflection coefficient as a function of frequency. If the function can be determined as a function of s then standard techniques may be used to obtain its poles and zeros and its analytic properties charted.

An important problem subset is where the impedance function $Z(s)$ is representable as a rational function of finite order in the numerator and denominator, $Z(s) = P(s)/Q(s)$. This is true if the impedance function represents the driving point impedance of a network of a finite number of passive resistors (R), capacitors (C) and inductors (L). The rational function representation is also frequently employed as an approximation in transmission line problems and in descriptions of metamaterials with a bounded upper frequency.

The function $Z(s)$ is positive real (PR) and hence there are a number of requirements on $P(s)$ and $Q(s)$. The literature on the topic is well developed and it is well known that,

1 As functions of s, the polynomials $P(s)$ and $Q(s)$ have real coefficients.

2 The order of $P(s)$ and $Q(s)$ can differ by at most 1.

3 The zeros of $P(s)$ and $Q(s)$ must occur only in the left half complex s-plane, possibly including the imaginary axis.

4 The real part of $Z(s)$ must be non-negative for all real frequencies ω.

Consequently, any algorithm that attempts to determine $Z(s)$ from real frequency data should ensure that all these conditions are met. Currently, there exist no well conditioned algorithms, valid for arbitrary data sets on arbitrary passive networks that can satisfy all of these requirements in a constructive manner. In fact, it may be impossible.

It is well known that A PR rational function can be synthesised using R, C and L elements and perfect transformers; the latter being unnecessary. In fact a rational function of finite order is PR if and only if it represents the driving point impedance of a network containing a finite number of R, L and C elements. Algorithms have been developed which synthesise an impedance function $Z(s)$ in terms of such components (e.g. [4]) which therefore guarantee the satisfaction of the requirements (1) to (4) above. Such methods can show good accuracy but LCR representations do not guarantee the uniqueness and ill-conditioning is a necessary consequence for some networks. Examples have been provided where LCR synthesis methods can be quite inaccurate [5].

This report is concerned with another frequency-domain method developed by Gustavsen, known as the vector fitting method [1]. Although the method does not guarantee positive realness (it does not employ requirement 4 as part of the construction) it has been found to be very accurate for quite a large class of networks and is quite popular and relatively simple to implement. Variants of the algorithm are introduced, one of which offers some significant advantages.

2.2 Frequency and impedance scaling and the reflection coefficient

In all variants of Gustavsen's vector fitting method proper frequency and impedance scaling is vital for numerical convergence and it is impossible to duplicate results without knowledge of the scale factor and sampling details. Scaling is important because the method variants are all non-linear and, in places, not always well conditioned.

If the impedance function is obtained from a complex reflection coefficient, all impedances are normalised with respect to the impedance of free space, but angle dependence is retained. Suppose a plane wave is reflected at an angle θ from a semi-infinite medium into free space. The complex reflection coefficient is then given by,

$$\Gamma = \frac{Z_{in}(\omega) - \overline{Z}_0}{Z_{in}(\omega) + \overline{Z}_0} \tag{2-1}$$

where $Z_{in}(\omega)$ is the input impedance in Ohms and $\overline{Z}_0$ is the reference impedance defined by,

$$\overline{Z}_0 = \begin{cases} Z_0/\cos\theta & \text{, for the TE polarisation} \\ Z_0\cos\theta & \text{, for the TM polarisation} \end{cases} \tag{2-2}$$

and $Z_0 \approx 377\Omega$ is the impedance of free space in Ohms. The impedance function is normalised such that $\hat{Z}(\omega) = Z_{in}(\omega)/Z_0$.

In examples where an input impedance $Z_{in}(\omega)$ is defined *ab-initio* in ohms, it is important to specify the reference impedance Z_0. This will be problem dependent.

Similarly, the frequency ω (and the poles and residues which take the same units) should also be dimensionless. For example, if the frequency f_{Hz} is specified in Hz, $\omega = f_{Hz}/f_{ref}$ where f_{ref} is also defined in Hz. The choice of f_{ref} is not obvious but numerical experiments suggest that over a uniformly sampled frequency interval best results are obtained when f_{ref} is the largest frequency on the interval.

2.3 An outline of Gustavsen's first and second variant

Gustavsen and Semlyen's first variant is described in [1]. There is a second variant described in [2] whose detailed formulation and implementation is rather more complicated. It is possible that an alternative and simpler formulation of this second variant exists, but no

details are available in [2]. Details of both variants are described in sections 2.7 and 2.8. Both variants are complicated by a requirement to ensure that conjugate poles and zeros are treated appropriately and differently to real poles and zeros and a requirement to ensure $Z(\omega) \rightarrow Z^{\star}(-\omega)$ for real ω, where the $\star$ represents the complex conjugate.

In outline, a general impedance function of rational form can be represented by,

$$Z(s) = \frac{P(s)}{Q(s)} = \sum_{m=1}^{N} \frac{r_m}{s - a_m} + d + s\,e \tag{2-3}$$

where there are N poles and the order of the polynomial $Q(s)$ is N. Poles exist is complex conjugate pairs or are real. If the complex pole pairs are grouped sequentially and ordered first on the m-list, then the complex poles are specified by complex values for a_m and r_m with $a_{2m-1} = a_{2m}^{\star}$ and $r_{2m-1} = r_{2m}^{\star}$. Real poles have real values for r_m and a_m. The coefficients d and e are real valued but need not always be present. If the order of $P(s)$ in $N-1$, then $d = e = 0$. If the order of $P(s)$ is N then $e = 0$ and $d \neq 0$. If the order of $P(s)$ is $N+1$ then both $e \neq 0$ and $d \neq 0$.

Suppose the impedance function is known at real frequencies and defined by $\hat{Z}(\omega)$ and that it is assumed to be approximately representable by $Z(s)$ so that $\hat{Z}(\omega) \approx Z(j\omega)$.

In the first variant, an attempt is made to identify the unknown parameters r_m, a_m d and e in (2-3) by assuming a representation,

$$\sigma(s)\hat{Z}(\omega) = p(s) \tag{2-4}$$

for all N_f frequencies ω_i, for $i = 1$ to N_f for which data is available. Here,

$$\sigma(s) = \sum_{m=1}^{N} \frac{\tilde{r}_m}{s - q_m} + 1 \tag{2-5}$$

and

$$p(s) = \sum_{m=1}^{N} \frac{r_m}{s - q_m} + d + se \tag{2-6}$$

where q_m are estimates of the pole positions which, in principle, can be arbitrary. The closer the pole estimates are to the correct positions the better behaved the algorithm so in practise there are conditions on their values as described later. For a given set of q_m, the left and right hand sides of (2-4) are linear in the unknown coefficients $\tilde{r}_m$, r_m, a_m d and e and (2-4) can be solved in a least squares sense to obtain these unknown coefficients. Once these are known, revised estimates for the poles may be determined as a function of $\tilde{r}_m$ and the previous estimates q_m. The algorithm may then be applied iteratively.

In general a weight function $w(\omega)$ is employed, such that $w(\omega_i)$ is real and positive for all ω_i and normalised such that,

$$\frac{1}{N_f} \sum_{i=1}^{N_f} w(\omega_i) = 1 \tag{2-7}$$

Now define a complex residual error vector with terms,

$$\varepsilon_i^{(1)} = w(\omega_i)\Big(\sigma(s_i)\hat{Z}(\omega_i) - p(s_i)\Big) \text{ , for } i \in [1, N_f] \tag{2-8}$$

and a positive error measure,

$$E_1 = \sum_{i=1}^{N_f} |\varepsilon_i^{(1)}|^2 \tag{2-9}$$

For diagnostic purposes the related RMS and relative error associated with E_1, may be defined by,

$$E_1^{(rms)} = \sqrt{\frac{1}{N_f}\sum_{i=1}^{N_f} |\varepsilon_i^{(1)}|^2} \tag{2-10}$$

and,

$$E_1^{(rel)} = \frac{E_1^{(rms)}}{\sqrt{\frac{1}{N_f}\sum_{i=1}^{N_f} |f(\omega_i)|^2}} \tag{2-11}$$

where $f(\omega) = w(\omega)\hat{Z}(\omega)$ is the weighted complex impedance.

In the absence of noise and least squares fitting errors, E_1 can be zero and no iteration is required. However, fitting errors can occur for a number of reasons due to numerical ill-conditioning of the least squares problem or the representation of the impedance function by a model of the wrong order. This can occur if N is unknown or infinite, if the order of $P(s)$ is not known or if the data contains noise. If multiple iterations are required, there is no guarantee of convergence [6] and several examples are provided later where this does not occur.

In Gustavsen's second variant [2], the form (2-5) is replaced by the more general representation,

$$\sigma(s) = \sum_{m=1}^{N} \frac{\tilde{r}_m}{s - q_m} + \tilde{d} \tag{2-12}$$

where $\tilde{d}$ is introduced to improve convergence. In this variation a second equation, linear in the unknown coefficients $\tilde{r}$ and $\tilde{d}$ is given by,

$$\Re\left\{\sum_{k=1}^{N_f}\left(\sum_{m=1}^{N}\frac{\tilde{r}_m}{s - q_m} + \tilde{d}\right)\right\} = N_f \tag{2-13}$$

is employed for normalisation. This results in the error measure E_1 being minimised subject to a different constraint on $\tilde{d}$. For a finite order model it is possible for both E_1 and E_2 to converge to zero, in which case $\sigma(s)$ must approach unity so that $\tilde{r}_m \to 0$ and $\tilde{d} \to 1$. The condition (2-13) implies that if $\tilde{r}_m \to 0$ then $\tilde{d} \to 1$ but does not ensure that such convergence occurs. Furthermore, E_2 can sometimes converge to a value that does not represent the global minimum and E_1 can be very small but represent a poor solution. The lack of a guarantee of proper convergence is true for all variants of the method in this report.

There is also a second error measure E_2 which should be considered, which does not include the $\tilde{r}_m$ or the $\tilde{d}$ terms. The associated error vector has terms,

$$\varepsilon_i^{(2)} = w(\omega_i)\Big(\hat{Z}(\omega_i) - p(s_i)\Big) \text{ , for } i \in [1, N_f] \tag{2-14}$$

and

$$E_2 = \sum_{i=1}^{N_f} |\varepsilon_i^{(2)}|^2 \tag{2-15}$$

As for E_1, the related RMS and relative error associated with E_2 are defined by,

$$E_2^{(rms)} = \sqrt{\frac{1}{N_f}\sum_{i=1}^{N_f} |\varepsilon_i^{(2)}|^2} \tag{2-16}$$

and,

$$E_2^{(rel)} = \frac{E_2^{(rms)}}{\sqrt{\frac{1}{N_f}\sum_{i=1}^{N_f} |f(\omega_i)|^2}} \tag{2-17}$$

Only if $\sigma(s) = 1$ will $E_1 = E_2$ and, as is shown later, this is not always the case. E_2 is not fully minimised in any of the algorithms reported here, but is the most direct measure of how close the impedance function is to the estimated pole-residue form.

2.4 Summary of new variants of the method

This section describes, in outline, some new variants of the method with full details of the algorithms provided later. The essence of the variants is to consider $\tilde{d}$ as a free variable and consider the set of variables $(\tilde{r_m}, r_m, \tilde{d}, d, e)$, as an eigenvector associated with the null eigenvalue of a matrix of terms similar to that employed in the least squares problem. The null eigenvalue method obtains a vector (ignoring in this summary the complication of real and complex residues),

$$\boldsymbol{x} = \alpha[\tilde{r}_1, \tilde{r}_2, ..., \tilde{r}_N; r_1, r_2, ..., r_N; \tilde{d}, d, e]^T \tag{2-18}$$

subject to the standard eigenvector normalisation $\boldsymbol{x}.\boldsymbol{x}^\star = 1$. The coefficient α is a positive constant to be determined and the vector $\boldsymbol{x}$ is real valued. This approach is not new and has been described in [11]. The new pole estimates may be determined directly from $\boldsymbol{x}$ and are independent of the normalisation but if the residues are required α is determined through a least squares procedure as described below. The vector $\boldsymbol{x}$, subject to the eigenvector normalisation, minimises E_1 when E_1 is sufficiently small. Alternatively, the $\tilde{d}$ term obtained from $\boldsymbol{x}$ may be used to re-minimise E_1 with the value of $\tilde{d}$ (possibly modified in some manner) assumed to be fixed.

The method for determination of $\boldsymbol{x}$ is well defined if there is only one null eigenvalue. However, if there is more than one or if there are several such eigenvalues close to zero,

minimisation of E_1 may not consistent with minimisation of E_2. This is because there exists a null space of possible eigenvectors which is not defined by a single eigenvector. E_1 will be zero for any member of the null space but most of the elements of this space will not minimise E_2. To resolve this difficulty, *all* the eigenvectors $\boldsymbol{x}_i$ which correspond to small eigenvalues are employed. Each vector $\boldsymbol{x}_i$ contains the tilde terms and the remaining terms $[r_1, r_2, ...r_N; d, e]$, considered as a sub-vector of $\boldsymbol{x}$. An unknown real coefficient, α_i, is associated with each of these eigenvectors and then the set of α_i on the sub-vectors is found which minimises E_2. The method is nice in that at no stage is it necessary to invert an ill-conditioned matrix. The null eigenvector space is well defined, even if individual null eigenvectors are non-unique, and the required number of null eigenvectors is much smaller than the total number of eigenvectors. The method has the feature that $\tilde{d}$ is not enforced to be unity when $\tilde{r}_m = 0$. Whether this is a good or a bad feature is unclear. In order to proceed in detail with the formulation, the single null eigenvector method is described first and subsequently generalised.

2.5 The single null-eigenvalue method

This section describes, in detail, the single null-eigenvalue method summarised above and described previously in [11]. The method is not particularly robust but serves as the foundation for the null-space method described later.

Suppose there are N poles, divided into P complex pairs and R real poles so that $N = 2P+R$. The complex poles are defined in terms of their real and imaginary parts by,

$$\tilde{q}_i = \tilde{q}'_i + j\tilde{q}''_i \text{ for } i = 1 \text{ to } P \tag{2-19}$$

with the remaining poles, for $i = 2P + 1$ to $2P + R$, real valued. The complex residues are similarly defined,

$$\tilde{r}_i = \tilde{r}'_i + j\tilde{r}''_i \text{ for } i = 1 \text{ to } P \tag{2-20}$$

and

$$r_i = r'_i + jr''_i \text{ for } i = 1 \text{ to } P \tag{2-21}$$

The remaining R real poles have real residues designated by $\tilde{r}_i$ and r_i for $i = 2P + 1$ to R. The unknown coefficients $\tilde{d}$, d and e are also real valued.

Note that although N is fixed through the iteration cycle, the numbers P and R will generally vary. In all current variants of the algorithms no constraints are placed to ensure a fixed value for P or R and it is not clear how this should be done if required.

Now consider a vector of unknown parameters defined, at each stage in the iteration cycle (where the pole estimates q_i for $i = 1$ to N are assumed given), by the vector of terms $\boldsymbol{x}$,

$$\boldsymbol{x} = [\tilde{\boldsymbol{r}}; \boldsymbol{r}; \tilde{d}; d, e]^T \tag{2-22}$$

where the sub-vectors,

$$\tilde{\boldsymbol{r}} = [(\tilde{r}'_1, \tilde{r}''_1), (\tilde{r}'_2, \tilde{r}''_2), \, ... \, , (\tilde{r}'_P, \tilde{r}''_P); \tilde{r}_{2P+1}, \tilde{r}_{2P+2}, \, ... \, , \tilde{r}_N]^T \tag{2-23}$$

$$\boldsymbol{r} = [(r'_1, r''_1), r'_2, r''_2), \, ... \, , (r'_P, r''_P); r_{2P+1}, r_{2P+2}, \, ... \, , r_N]^T \tag{2-24}$$

The vector of terms $\boldsymbol{x}$ is of dimension N_x where N_x is of minimum dimension $2N+1$ when there are no d or e terms and maximum dimension $N_x = 2N+3$ when both d and e terms are present. The intermediate case is when there is is no e term in which case $N_x = 2N+2$.

If it is known *a-priori* that the the impedance function $Z(s) \to \infty$ as $s \to \infty$ then both e and d terms are required. If it is known that $Z(s) \to 0$ as $s \to \infty$ then it is preferable to assume neither e and d terms are present. If $Z(s)$ remains finite as $s \to \infty$ and when $s \to 0$ then it is preferable to assume the e term is absent. In practise, for an unknown impedance function it does little harm to assume both e and d terms are required and rely on these to be assigned near zero values if this provides a best least squares fit. This is particularly true for transmission line problems where the true model order is infinite and the impedance at infinity is undefined.

Now associate a rectangular matrix A, of dimension $N_f \times N_x$. The matrix is composed of row vectors $\boldsymbol{A}_i$,

$$A = \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \\ .. \\ .. \\ \boldsymbol{A}_{N_f} \end{pmatrix} \tag{2-25}$$

where,

$$\boldsymbol{A}_i = [\boldsymbol{A}_i^{(1)}; \boldsymbol{A}_i^{(2)}; w(\omega_i)\hat{Z}(\omega_i); -w(\omega_i), -w(\omega_i)s_i] \tag{2-26}$$

The final two terms to the row vectors are both present only when $N_x = 2N+3$, otherwise one or both of these terms is omitted from the definition (both terms omitted when $N_x = 2N+1$ and the last term omitted when $N_x = 2N+2$). The sub-vectors $\boldsymbol{A}_i^{(1)}$ and $\boldsymbol{A}_i^{(2)}$ are defined by,

$$\begin{aligned} \boldsymbol{A}_i^{(1)} = w(\omega_i)\hat{Z}(\omega_i) \Bigg[& \left(\frac{1}{s_i - q_1} + \frac{1}{s_i - q_1^\star} \, , \, \frac{j}{s_i - q_1} - \frac{j}{s_i - q_1^\star} \right) , \\ & \left(\frac{1}{s_i - q_2} + \frac{1}{s_i - q_2^\star} \, , \, \frac{j}{s_i - q_2} - \frac{j}{s_i - q_2^\star} \right) , \, ... \, , \left(\frac{1}{s_i - q_P} + \frac{1}{s_i - q_P^\star} \, , \, \frac{j}{s_i - q_P} - \frac{j}{s_i - q_P^\star} \right) ; \\ & \frac{1}{s_i - q_{2P+1}} \, , \, \frac{1}{s_i - q_{2P+2}} \, , \, ... \, , \, \frac{1}{s_i - q_{2P+R}} \Bigg] \end{aligned} \tag{2-27}$$

and

$$\boldsymbol{A}_i^{(2)} = -w(\omega_i)\left[\left(\frac{1}{s_i-q_1}+\frac{1}{s_i-q_1^\star}\,,\,\frac{j}{s_i-q_1}-\frac{j}{s_i-q_1^\star}\right),\right.$$
$$\left(\frac{1}{s_i-q_2}+\frac{1}{s_i-q_2^\star}\,,\,\frac{j}{s_i-q_2}-\frac{j}{s_i-q_2^\star}\right),\ \ldots\ ,\left(\frac{1}{s_i-q_P}+\frac{1}{s_i-q_P^\star}\,,\,\frac{j}{s_i-q_P}-\frac{j}{s_i-q_P^\star}\right);$$
$$\left.\frac{1}{s_i-q_{2P+1}}\,,\,\frac{1}{s_i-q_{2P+2}}\,,\,\ldots\,,\frac{1}{s_i-q_{2P+R}}\right] \tag{2-28}$$

With these definitions, $A\boldsymbol{x} \approx 0$. More precisely, $A\boldsymbol{x} = \boldsymbol{\varepsilon}^{(1)}$ where,

$$\boldsymbol{\varepsilon}^{(1)} = [\varepsilon_1^{(1)}, \varepsilon_2^{(1)}, \ \ldots, \varepsilon_{N_f}^{(1)}]^T \tag{2-29}$$

In order to find the least squares solution for $\boldsymbol{x}$, the Matrix A is partitioned into real and imaginary parts,

$$A = A' + jA'' \tag{2-30}$$

and similarly for the residual vector $\boldsymbol{\varepsilon}^{(3)}$,

$$\boldsymbol{\varepsilon}^{(1)} = \boldsymbol{\varepsilon}^{(1)\prime} + j\boldsymbol{\varepsilon}^{(1)\prime\prime} \tag{2-31}$$

with residual error E_1, defined in (2-9). Consequently,

$$\left[\begin{array}{c} A' \\ A'' \end{array}\right]\boldsymbol{x} = \left[\begin{array}{c} \boldsymbol{\varepsilon}^{(1)\prime} \\ \boldsymbol{\varepsilon}^{(1)\prime\prime} \end{array}\right] \tag{2-32}$$

where the partitioned matrix on the left hand side is referred to as $\hat{A}$. If the residual error E_1 can be made zero for some choice of $\tilde{d}$, then $\boldsymbol{x}$ may be obtained as a vector proportional to the eigenvector corresponding to the null eigenvalue, $\lambda = 0$, of

$$B\hat{\boldsymbol{x}} = \lambda\hat{\boldsymbol{x}} \tag{2-33}$$

such that

$$\boldsymbol{x} = \alpha\hat{\boldsymbol{x}}_0 \tag{2-34}$$

where

$$B = \hat{A}^T\hat{A} = A'^T A' + A''^T A'' \tag{2-35}$$

and α is some positive real constant yet to be determined. The T represents the transpose since $\hat{A}$ is real valued. Thus B is real symmetric and all the eigenvectors and eigenvalues are real valued as required.

When the residual error E_1 can not be made zero, the required $\hat{\boldsymbol{x}}_0$ is taken numerically as the eigenvector of (2-33) associated with the eigenvalue λ whose magnitude $|\lambda|$ is smallest. Provided E_1 is sufficiently small, this provides a least squares solution for $\boldsymbol{x}$ that minimises E_1

subject to the eigenvector normalisation $\hat{\boldsymbol{x}}_0.\hat{\boldsymbol{x}}_0 = 1$. The eigenvalue problem is well defined and standard library routines exist for this purpose. Here the LaPack routine DSYEV is employed, for a real symmetric matrix. If there is only one eigenvalue whose magnitude is much smaller than all the others, the solution is unique and properly defined. If there are several eigenvalues whose magnitude is small, the algorithm should be modified as discussed in the next section.

Once $\hat{\boldsymbol{x}}_0$ is computed, new estimates may be determined for the pole positions q_i which are independent of α. In order to determine the residue coefficients, α may be determined as the coefficient that minimises the second residual, E_2. It may be shown (see appendix A, Case 1), that

$$\alpha = \frac{-\Re(\boldsymbol{f}^\star.\boldsymbol{c})}{\boldsymbol{c}^\star.\boldsymbol{c}} \tag{2-36}$$

where $\boldsymbol{f}$ is the complex vector of terms, $f(\omega_i) = w(\omega_i)\hat{Z}(\omega_i)$, defined by

$$\boldsymbol{f} = [w(\omega_1)\hat{Z}(\omega_1), w(\omega_2)\hat{Z}(\omega_2), \ ... \ , w(\omega_{N_f})\hat{Z}(\omega_{N_f})]^T \tag{2-37}$$

and

$$\boldsymbol{c} = A_s\hat{\boldsymbol{x}}^{(s)} \tag{2-38}$$

where A_s is the matrix of dimension $N_f \times N_2$, where $N_2 = N_x - N - 1$, given by,

$$A_s = \begin{pmatrix} \boldsymbol{A}_1^{(s)} \\ \boldsymbol{A}_2^{(s)} \\ .. \\ .. \\ \boldsymbol{A}_{N_f}^{(s)} \end{pmatrix} \tag{2-39}$$

where the complex row vectors $\boldsymbol{A}_i^{(s)}$ are given by,

$$\boldsymbol{A}_i^{(s)} = [\boldsymbol{A}_i^{(2)}; -w(\omega_i), -w(\omega_i)s_i] \tag{2-40}$$

and the real column vector $\hat{\boldsymbol{x}}^{(s)}$ is given by the terms $N+1$ to N_x, excluding the term $2N+1$, of $\hat{\boldsymbol{x}}_{\boldsymbol{0}}$ corresponding to the sub-vectors $\boldsymbol{r}$, d and e. In particular, $\hat{\boldsymbol{x}}_{\boldsymbol{0}}$ may be written as,

$$\hat{\boldsymbol{x}}_0 = [\boldsymbol{x}_0^{(u)}; \boldsymbol{x}_0^{(v)}; x_0^{(\tilde{d})}, x_0^{(d)}, x_0^{(e)}]^T \tag{2-41}$$

where $\boldsymbol{x}_0^{(u)}$ is of dimension N and represents the unscaled $\boldsymbol{r}$ vector, $\boldsymbol{x}_0^{(v)}$ is of dimension N and represents the unscaled $\tilde{\boldsymbol{r}}$ vector and $x_0^{(\tilde{d})}, x_0^{(d)}$ and $x_0^{(e)}$ represent the unscaled terms for $\tilde{d}, d$ and e respectively. So,

$$\hat{\boldsymbol{x}}^{(s)} = [\boldsymbol{x}_0^{(v)}; x_0^{(d)}, x_0^{(e)}]^T \tag{2-42}$$

with the understanding that the last two terms may not be present.

Once α is computed, $\tilde{d}$ is given by,

$$\tilde{d} = \alpha x_0^{(d)} \tag{2-43}$$

and

$$\tilde{\boldsymbol{r}} = \alpha \boldsymbol{x_0^{(u)}} \tag{2-44}$$

Actually, in the single null-eigenvalue method there is no need to compute α until the process has completed since the iteration process only requires updates of the pole positions q_i and these are independent of α (though it is needed to monitor E_2 to check for convergence).

The pole positions may be determined from $\boldsymbol{x}_0^{(u)}$ and $x_0^{(\tilde{d})}$ (with $\tilde{\boldsymbol{r}}$ replaced by $\boldsymbol{x}_0^{(u)}$ and $\tilde{d}$ replaced by $x_0^{(\tilde{d})}$) as the eigenvalues of the matrix,

$$H = H^{(1)} - \frac{1}{\tilde{d}} H^{(2)} \tag{2-45}$$

as defined in appendix B. The Lapack routine DGEEV (for determining the eigenvalues of a real non-symmetric matrix) is employed for this purpose. As in [1], if a pole is estimated to have positive real part (which is unphysical for a passive impedance function) its real part is flipped (made negative).

The previous values of q_i and the previous value of P (describing how many of the poles are complex), are updated and the process begins again until convergence is achieved or the number of iterations reaches a pre-set maximum. It is important to note that there is no proof that convergence will occur and observations of the method show that the error measure E_2 can vary erratically, in a highly non-monotonic manner, as iteration proceeds.

2.6 The null-space method

The null-space method is a more robust generalisation of the single null eigenvalue method. As there, the matrix B as defined in (2-35) is constructed and the eigenvalue equation (2-33) is solved. Now, however, instead of selecting just the single eigenvector $\hat{\boldsymbol{x}}_\mathbf{0}$ corresponding to the smallest singular value (eigenvalue magnitude) for $|\lambda|$, the set of N_s eigenvectors $\hat{\boldsymbol{x}}_\boldsymbol{i}$ is selected which correspond to singular values below a pre-defined threshold. Eigenvectors are selected whose singular values meets one of two criteria;

(1) If the singular value is less than $\epsilon_m |\lambda|_{max}$, where ϵ_m is the machine precision (approximately 2×10^{-16} in Double Precision arithmetic on a 64 bit PC) and $|\lambda|_{max}$ is the maximum singular value.
(2) If the singular value is less than $L|\lambda|_{min}$ where $|\lambda|_{min}$ is the minimum singular value and L is a fairly large number. A typical value of $L = 1000$ is generally employed, but is subject to user control.

There is also an additional requirement, $N_s \leq N_2$, or the vector $\boldsymbol{\alpha}$ is not unique. This places an upper bound on the number of eigenvectors selected by thresholding on the singular values.

These eigenvectors are employed to construct the matrix X_r as defined in case 2 of appendix A in order to find that linear combination of selected eigenvectors that minimises E_2. The coefficients α_i that define the linear combination are then used to obtain the residues $\tilde{r}_i$ and constant $\tilde{d}$. The requirement for a threshold parameter in the definition of the generalised inverse in Gustavsen's first and second variants is replaced, in this method, by a selection requirement to decide which eigenvectors are to be used in defining the null-space.

Examples will be given showing how this method compares with Gustavsen's first and second variant.

2.7 Gustavsen's first and a modification of Gustavsen's second variant

The null-eigenvalue method may also be used to provide an alternative method for obtaining $\tilde{d}$ instead of employing the normalisation in (2-12) in Gustavsen's second method. Here (2-43) is employed to determine $\tilde{d}$, though some modifications are needed due to the lack of robustness of the method. The method is used to estimate $\tilde{d}$ when $\tilde{d}$ is predicted to be fairly small and positive, which is where the second variant was demonstrated to obtain superior convergence. Under the prescription,

(1) If $\tilde{d}$ is predicted to lie between $\tilde{d}_{min}$ and 1 then it is unchanged, where $\tilde{d}_{min}$ is a small positive number (a value $\tilde{d}_{min} = 10^{-6}$ is assumed).
(2) If $\tilde{d}$ is predicted to lie between $-\tilde{d}_{min}$ and $\tilde{d}_{min}$ redefine $\tilde{d} = \tilde{d}_{min}$
(3) If $\tilde{d}$ is predicted less than $-\tilde{d}_{min}$ or greater than $\tilde{d}_{max}$ it is reassigned the value $\tilde{d} = 1$, where $\tilde{d}_{max}$ is some positive number not much smaller than 1 (here, $\tilde{d}_{max} = 0.5$).

The prescription reverts to Gustavsen's first variant [1], for which $\tilde{d} = 1$, if the predicted value of $\tilde{d}$ by the null-eigenvalue method is outside of the preset bounds. Although this prescription seems somewhat arbitrary it appears to work quite well in problems for which the model order is not overestimated. Once $\tilde{d}$ is defined, the method proceeds as described below. The algorithm is similar to the first variant with minor modifications, in contrast with the second variant for which $\tilde{d}$ is a function of the residues of $\sigma(s)$.

First, define the matrix A_g composed of row vectors $\boldsymbol{A}_i^{(g)}$,

$$A_g = \begin{pmatrix} \boldsymbol{A}_1^{(g)} \\ \boldsymbol{A}_2^{(g)} \\ .. \\ .. \\ \boldsymbol{A}_{N_f}^{(g)} \end{pmatrix} \tag{2-46}$$

where,

$$\boldsymbol{A}_i^{(g)} = [\boldsymbol{A}_i^{(1)}; \boldsymbol{A}_i^{(2)}; -w(\omega_i), -w(\omega_i)s_i] \tag{2-47}$$

These definitions are similar to those for the matrix A composed of row vectors $\boldsymbol{A}_i$ in (2-25) and (2-26) but omit the $w(\omega_i)\hat{Z}(\omega_i)$ terms. The least squares solution for $\boldsymbol{x}_g$ is then obtained as,

$$A_g \boldsymbol{x}_g = -\tilde{d}\boldsymbol{f} + \boldsymbol{\varepsilon}^{(1)} \tag{2-48}$$

that minimises E_1. The vector of values $\boldsymbol{x}_g$ takes the form,

$$\boldsymbol{x}_g = [\tilde{\boldsymbol{r}}; \boldsymbol{r}; d, e]^T \tag{2-49}$$

similar to the definition for $\boldsymbol{x}$ in (2-22) with the absence of the $\tilde{d}$ term. Since A_g and $\boldsymbol{f}$ are complex and $\boldsymbol{x}_g$ is real, A_g and $\boldsymbol{f}$ may be expressed in terms of their real and imaginary parts,

$$\begin{aligned} A_g &= A'_g + jA''_g \\ \boldsymbol{f} &= \boldsymbol{f}' + j\boldsymbol{f}'' \end{aligned} \tag{2-50}$$

leading to the least-squares equation,

$$(A_g'^T A'_g + A_g''^T A''_g)\boldsymbol{x}_g = -\tilde{d}(A_g'^T \boldsymbol{f}' + A_g''^T \boldsymbol{f}'') \tag{2-51}$$

where the T superfix represents the matrix transpose. The bracketed quantity on the left is a square matrix of dimension $N_x - 1$ which may sometimes be invertible but is generally ill-conditioned. The equation (2-51) has the equivalent form,

$$\Re(A_g^\dagger A_g)\boldsymbol{x}_g = -\tilde{d}\,\Re(A_g^\dagger \boldsymbol{f}) \tag{2-52}$$

where the $\dagger$ superfix represents the matrix complex conjugate transpose and $\Re(x)$ represents the real part of x for any x scalar, vector or matrix. In the latter representation, the formal solution is given by,

$$\boldsymbol{x}_g = -\tilde{d}\,Q^{-1}\,\Re(A_g^\dagger \boldsymbol{f}) \tag{2-53}$$

where Q^{-1} is the Moore-Penrose inverse of Q. This may be defined using a singular value decomposition (SVD), such that

$$Q = \Re(A_g^\dagger A_g) \equiv U\Sigma V^T \tag{2-54}$$

where Σ is the diagonal matrix of singular values (positive real) taken in descending order and U and V are left and right real orthogonal matrices. When the singular values fall below a specified threshold, $s_{min} > 0$, they are defined to be zero so that,

$$\Sigma = \begin{pmatrix} \sigma_1 & & & & & & \\ & \sigma_2 & & & & & \\ & & . & & & & \\ & & & . & & & \\ & & & & . & & \\ & & & & & \sigma_s & \\ & & & & & & 0 \\ & & & & & & & 0 \end{pmatrix} \tag{2-55}$$

for $\sigma_1 = \sigma_{max} \geq \sigma_2 \geq \ ... \ \geq \sigma_s > s_{min}$. The Moore-Penrose pseudo-inverse is then given by,

$$Q^{-1} = (\Re(A_g^\dagger A_g))^{-1} = V\Sigma^{-1}U^T \tag{2-56}$$

where

$$\Sigma^{-1} = \begin{pmatrix} 1/\sigma_1 & & & & & & & \\ & 1/\sigma_2 & & & & & & \\ & & . & & & & & \\ & & & . & & & & \\ & & & & . & & & \\ & & & & & 1/\sigma_s & & \\ & & & & & & 0 & \\ & & & & & & & 0 \end{pmatrix} \tag{2-57}$$

As previously mentioned, the value for s_{min} can have a substantial impact on the accuracy and convergence of the algorithm when Q is almost singular. For best results, s_{min} should be defined by,

$$s_{min} = \epsilon_D \sigma_{max} \tag{2-58}$$

where ϵ_D is the machine precision ($\approx 2 \times 10^{-16}$ in double precision arithmetic on a 64 bit computer). Numerical experiments suggest that numbers significantly bigger or smaller than this appear to be sub-optimal. The LaPack routine DGELSS is employed to obtain the SVD minimum norm solution for $\boldsymbol{x}_g$.

Once $\boldsymbol{x}_g$ is computed, and consequently the residues $\tilde{\boldsymbol{r}}$ for $\sigma(s)$ are defined, new estimates of the pole positions q_i are obtained as the eigenvalues of the matrix H, as defined in appendix B.

2.8 Gustavsen's second variant

In [2] no details were presented showing how the linear equations were solved with the given normalisation. Here a formulation is provided. The algorithm appears quite complicated and it is not known if there is a special simplification or whether general purpose linear constraint algorithms were applied in the original paper.

The first equation is given by (2-48), described above. The second equation is given by (2-13). In order to identify terms with subsets of the matrix A, and to use an arbitrary weight function $w(\omega)$, (2-13) is replaced by its generalisation,

$$\Re\left\{\sum_{k=1}^{N_f} w(\omega_k)\left(\sum_{m=1}^{N}\frac{\tilde{r}_m}{s-q_m} + \tilde{d}\right)\right\} = N_f \tag{2-59}$$

This assumes the normalisation of the weight function defined in (2-7). Equation (2-59) may be written in matrix notation using subsets of the matrix A and vector $\boldsymbol{x}$. For this purpose

define,

$$A_t = \begin{pmatrix} \boldsymbol{A}_1^{(t)} \\ \boldsymbol{A}_2^{(t)} \\ .. \\ .. \\ \boldsymbol{A}_{N_f}^{(t)} \end{pmatrix} \tag{2-60}$$

where the row vectors,

$$\boldsymbol{A}_i^{(t)} = [\boldsymbol{A}_i^{(1)}] \tag{2-61}$$

A_t is of dimension $N_f \times N$. Similarly, define

$$\boldsymbol{x}_t = [\tilde{\boldsymbol{r}}]^T \tag{2-62}$$

of dimension N. In addition, define $\mathbf{1}_L = [1, 1, \ldots, 1]$ to be the row vector of ones of dimension L. So (2-59) may be re-written as,

$$\Re(\mathbf{1}_{N_f} A_t)\boldsymbol{x}_t + N_f \tilde{d} = N_f \tag{2-63}$$

so that,

$$\tilde{d} = 1 - \frac{1}{N_f}\Re(\mathbf{1}_{N_f} A_t)\boldsymbol{x}_t \tag{2-64}$$

Combining (2-64) and (2-48),

$$\boldsymbol{\varepsilon}^{(1)} = A_g \boldsymbol{x}_g - \frac{1}{N_f}\Re(\mathbf{1}_{N_f} A_t)\boldsymbol{x}_t \boldsymbol{f} + \boldsymbol{f} \tag{2-65}$$

The error E_1 is now minimised with respect to the terms in the vector $\boldsymbol{x}_g$ noting that $\boldsymbol{x}_t \subset \boldsymbol{x}_g$. To do this, note that $\boldsymbol{x}_g = [\boldsymbol{x}_t; \boldsymbol{x}_s]$ where $\boldsymbol{x}_s = [\boldsymbol{r}; d, e]$, assuming the d and e terms are present. In general $\boldsymbol{x}_s$ is of dimension N_2, recalling that $N_2 = N$ if neither d nor e terms are present, $N_2 = N + 1$ if only the d term is present or $N_2 = N + 2$ if both additional terms are present. The matrix A_g is similarly partitioned with $A_g = [A_t; A_s]$, where

$$A_s = \begin{pmatrix} \boldsymbol{A}_1^{(s)} \\ \boldsymbol{A}_2^{(s)} \\ .. \\ .. \\ \boldsymbol{A}_{N_f}^{(s)} \end{pmatrix} \tag{2-66}$$

with the row vectors (again assuming the d and e terms are present),

$$\boldsymbol{A}_i^{(s)} = [\boldsymbol{A}_i^{(2)}; -w(\omega_i), -w(\omega_i)\, s_i] \tag{2-67}$$

So,

$$A_g \boldsymbol{x}_g = A_t \boldsymbol{x}_t + A_s \boldsymbol{x}_s \tag{2-68}$$

Let the terms in the matrices A_t and A_s be signified as $A_{kj}^{(t)}$ and $A_{kj}^{(s)}$, respectively. Similarly let the terms in the vectors $\boldsymbol{x}_t$ and $\boldsymbol{x}_s$ be signified as $x_j^{(t)}$ and $x_j^{(s)}$, respectively. Substituting (2-68) in (2-65) and rewriting in index notation using the Einstein summation convention,

$$\varepsilon_k^{(1)} = A_{kj}^{(s)} x_j^{(s)} + A_{kj}^{(t)} x_j^{(t)} - \frac{1}{N_f} f_k \, \Re(A_{lj}^{(t)} 1_l) x_j^{(t)} + f_k \tag{2-69}$$

So,

$$E_1 = u_1 + u_2 + u_2^\star \tag{2-70}$$

where

$$u_1 = A_{kj}^{(s)\star} A_{ki}^{(s)} x_j^{(s)} x_i^{(s)} + A_{kj}^{(t)\star} A_{ki}^{(t)} x_j^{(t)} x_i^{(t)} + \frac{1}{N_f^2} f_k^\star f_k \Re(A_{lj}^{(t)} 1_l) \, \Re(A_{mi}^{(t)} 1_m) \, x_i^{(t)} x_j^{(t)} + f_k^\star f_k \tag{2-71}$$

and

$$\begin{aligned} u_2 =& A_{kj}^{(s)\star} A_{ki}^{(t)} x_i^{(t)} x_j^{(s)} - \frac{1}{N_f} A_{kj}^{(s)\star} x_j^{(s)} f_k \Re(A_{li}^{(t)} 1_l) x_i^{(t)} \\ &+ A_{kj}^{(s)\star} x_j^{(s)} f_k - \frac{1}{N_f} A_{kj}^{(t)\star} x_j^{(t)} f_k \Re(A_{li}^{(t)} 1_l) x_i^{(t)} \\ &+ A_{kj}^{(t)\star} x_j^{(t)} f_k - \frac{1}{N_f} f_k^\star f_k \, \Re(A_{li}^{(t)} 1_l) x_i^{(t)} \end{aligned} \tag{2-72}$$

The error is minimised when both $\partial E_1 / \partial x_p^{(s)} = 0$ and $\partial E_1 / \partial x_p^{(t)} = 0$. This leads to the two equations,

$$\Re(A_{kp}^{(s)\star} A_{ki}^{(s)}) x_i^{(s)} + \Re(A_{kp}^{(s)\star} A_{ki}^{(t)}) x_i^{(t)} - \frac{1}{N_f} \Re(A_{kp}^{(s)\star} f_k) \, \Re(A_{li}^{(t)} 1_l) x_i^{(t)} = -\Re(A_{kp}^{(s)\star} f_k) \tag{2-73}$$

and

$$\begin{aligned} &\Re(A_{kp}^{(t)\star} A_{ki}^{(t)}) x_i^{(t)} + \frac{1}{N_f^2} f_k^\star f_k \Re(A_{lp}^{(t)} 1_l) \, \Re(A_{mi}^{(t)} 1_m) x_i^{(t)} \\ &+ \Re(A_{kj}^{(s)\star} A_{kp}^{(t)}) x_j^{(s)} - \frac{1}{N_f} \Re(A_{kj}^{(s)\star} f_k) \, \Re(A_{lp}^{(t)} 1_l) x_j^{(s)} \\ &- \frac{1}{N_f} \Re(A_{kp}^{(t)\star} f_k) \, \Re(A_{li}^{(t)} 1_l) x_i^{(t)} - \frac{1}{N_f} \Re(A_{kj}^{(t)\star} f_k) \, \Re(A_{lp}^{(t)} 1_l) x_j^{(t)} \\ &= -\Re(A_{kp}^{(t)\star} f_k) + \frac{1}{N_f} \Re(f_k^\star f_k) \, \Re(A_{lp}^{(t)} 1_l) \end{aligned} \tag{2-74}$$

where $f_k^\star f_k$ is real so the $\Re$ operator is superfluous on this term. Equations (2-73) and (2-74) can be written in matrix notation as,

$$\begin{aligned} S_1 \boldsymbol{x}_s + T_1 \boldsymbol{x}_t &= \boldsymbol{F}_1 \\ S_2 \boldsymbol{x}_s + T_2 \boldsymbol{x}_t &= \boldsymbol{F}_2 \end{aligned} \tag{2-75}$$

where S_1 is a real square matrix of dimension $N_2 \times N_2$ and T_1 is generally a real rectangular matrix of dimension $N_2 \times N$. S_2 is real rectangular of dimension $N \times N_2$ and T_2 is real square of dimension $N \times N$. These are given by,

$$S_1 = \Re(A_s^\dagger A_s) \tag{2-76}$$

$$T_1 = \Re(A_s^\dagger A_t) - \frac{1}{N_f}\boldsymbol{\tau}_1^T\boldsymbol{\tau}_2 \tag{2-77}$$

$$S_2 = \Re(A_t^\dagger A_s) - \frac{1}{N_f}\boldsymbol{\tau}_3^T\boldsymbol{\tau}_4 \tag{2-78}$$

and

$$T_2 = \Re(A_t^\dagger A_t) - \frac{1}{N_f}(\boldsymbol{\tau}_5^T\boldsymbol{\tau}_2 + \boldsymbol{\tau}_2^T\boldsymbol{\tau}_5) + \frac{1}{N_f^2}(\boldsymbol{f}^\dagger\boldsymbol{f})\boldsymbol{\tau}_2^T\boldsymbol{\tau}_2 \tag{2-79}$$

The vectors $\boldsymbol{F}_1$ and $\boldsymbol{F}_2$ are given by,

$$\boldsymbol{F}_1 = -\boldsymbol{\tau}_1^T \tag{2-80}$$

$$\boldsymbol{F}_2 = -\boldsymbol{\tau}_5^T + \frac{1}{N_f}(\boldsymbol{f}^\dagger\boldsymbol{f})\,\boldsymbol{\tau}_2^T \tag{2-81}$$

of dimension N_2 and N, respectively. The $\boldsymbol{\tau}$ vectors are given by,

$$\begin{array}{ll}
\boldsymbol{\tau}_1^T = \Re(A_s^\dagger\boldsymbol{f}) & \text{of dimension } N_2 \\
\boldsymbol{\tau}_2 = \mathbf{1}_{N_f}\,\Re(A_t) & \text{of dimension } N \\
\boldsymbol{\tau}_3^T = \Re(A_t^\dagger)\,\mathbf{1}_{N_f}^T = \boldsymbol{\tau}_2^T & \text{of dimension } N \\
\boldsymbol{\tau}_4 = \Re(\boldsymbol{f}^\dagger A_s) & \text{of dimension } N_2 \\
\boldsymbol{\tau}_5^T = \Re(A_t^\dagger\boldsymbol{f}) & \text{of dimension } N
\end{array} \tag{2-82}$$

The two equations (2-75) may be combined as a single matrix equation,

$$\Omega\boldsymbol{x}_g = \boldsymbol{F} \tag{2-83}$$

where Ω is the $(N + N_2) \times (N + N_2)$ real square matrix,

$$\Omega = \left(\begin{array}{c|c} T_1 & S_1 \\ \hline T_2 & S_2 \end{array}\right) \tag{2-84}$$

and

$$\boldsymbol{F} = [\boldsymbol{F}_1; \boldsymbol{F}_2]^T \tag{2-85}$$

Equation (2-83) may be solved as $\boldsymbol{x}_g = \Omega^{-1}\boldsymbol{F}$ where Ω^{-1} is the Moore-Penrose inverse as defined in Gustavsen's first method variant, described above. As there, it is assumed that $s_{min} = \epsilon_D\sigma_{max}$ where ϵ_D is the machine precision.

This method, where the residuals are determined directly from $\boldsymbol{x}_g$, will be referred to as **Gustavsen's variant 2a**. However, this method can be employed just to determine the coefficient $\tilde{d}$, using the predicted $\boldsymbol{x}_t$ sub-vector from $\boldsymbol{x}_g$, as defined in (2-64). Once $\tilde{d}$ is determined the poles and residues can be predicted using the method described in section 2.7, instead of employing the values predicted directly in $\boldsymbol{x}_g$. Here this is referred to as **Gustavsen's variant 2b**. The $\tilde{d}$ employed in variant 2b does *not* satisfy (2-59) but instead permits the optimal choice of $\tilde{r}_i$ for the given $\tilde{d}$. I.e. E_1 is minimised subject to the given $\tilde{d}$. The variants 2a and 2b are different and produce different results, as will be demonstrated later.

In either of the two variants, the vector of residue coefficients $\boldsymbol{x}_g$ that minimises E_1 subject to the constraint (2-59) only exists if the vector norm[1] $||\Omega\boldsymbol{x}_g - \boldsymbol{F}||^2 = 0$. This is true if Ω possesses a proper inverse but frequently this is not the case in which case the Moore-Penrose is used as previously defined. When the proper inverse does not exist, either there are multiple (infinitely many) $\boldsymbol{x}_g$ that satisfy $||\Omega\boldsymbol{x}_g - \boldsymbol{F}||^2 = 0$ (the Moore-Penrose inverse defining that solution for which $||\boldsymbol{x}_g||^2$ is smallest) or there is no solution for which $||\Omega\boldsymbol{x}_g - \boldsymbol{F}||^2 = 0$ (in which case the Moore-Penrose inverse provides the solution which minimises $||\Omega\boldsymbol{x}_g - \boldsymbol{F}||^2$). There are some instances for which there is no numerical solution and $||\Omega\boldsymbol{x}_g - \boldsymbol{F}||^2$ is quite large when employing the previously defined algorithm. This can occur for problems where the impedance function is finite rational and our choice of N is greater than the number of poles in the function. Physically this tells us that our estimates of pole positions is sufficiently poor that it is not possible to meet the constraint (2-59) assuming finite residues for which there is a small margin of error (defined by numerical precision). However, such events are rare and most of the time $||\Omega\boldsymbol{x}_g - \boldsymbol{F}||^2$ is small but non zero in such circumstances.

2.9 A note on the error measures in the different methods

In Gustavsen's first variant the measure E_1 is minimised at each iteration subject to $\tilde{d} = 1$. There is a theoretical possibility that E_1 can not be numerically minimised if the matrix Q (see (2-54)) does not possess an inverse. This can occur if the Moore-Penrose solution results in the vector norm $||\Re(A_g^\dagger A_g)\boldsymbol{x}_g + \tilde{d}\Re(A_g^\dagger \boldsymbol{f})||^2 > \delta$, where δ is a very small number, but (so far) no examples of this pathology have been observed. The requirement that $\tilde{d} = 1$ is a condition for a fully converged solution but is not optimal for initial iterations. No use is made of the E_2 measure.

In Gustavsen's second variant, and our derivatives of it, E_1 is minimised subject to an alternative constraint (2-59) on $\tilde{d}$. This can accelerate the iteration process (2-12) but, as some of the later exampled demonstrate, can be less robust. There are occasions for which it is not possible to numerically minimise E_1 subject to this constraint as described above. When $||\Omega\boldsymbol{x}_g - \boldsymbol{F}||^2 \neq 0$ E_1 is not minimised and there can be significant differences employing

[1] The vector norm $||\boldsymbol{x}||^2$ of the vector $\boldsymbol{x}$ is defined by the dot-product $\boldsymbol{x}^\star.\boldsymbol{x}$ where the $\star$ represents the complex conjugate.

variant 2a or 2b. In the latter, the fact that $||\Omega \boldsymbol{x}_g - \boldsymbol{F}||^2 \neq 0$ is not so important since this affects only the choice of $\tilde{d}$ in a second minimisation of E_1 as employed in Gustavsen's first variant. Again, no use is made of the E_2 measure.

In the null-space method, the measure E_1 is minimised when E_1 can be made small enough, i.e. where there exists a linear combination of variables for which E_1 is close to zero. A null-space of eigenvectors is then selected that approximately minimises E_1 and minimises E_2. This is a partial minimisation of E_2 from those eigenvectors representing the null-space. If E_1 cannot be made small enough, then this will result in multiple eigenvectors associated with singular values that are not sufficiently small, indicating a poor model. This may result in failure of convergence, but under these conditions the other variants may also fail to converge. The method appears to be a compromise of sorts, where residue coefficients are chosen to reduce both E_1 and E_2. Note that the thresholding parameter L should not be chosen so large that that the eigenvectors are associated with large singular values, since this looses the ability to reduce E_1 and poor convergence may be expected. Currently there exists no satisfactory analysis of convergence of any variants of the method.

3 Numerical Examples

3.1 Summary

In this section different variants are compared and contrasted using a number of examples. Some of these come from [1] and [2] though sampling information was lacking in these papers. The final example considers a practical application for the study of the angle dependent input impedance, generated from a complex reflection coefficient, of a composite radar absorbing material. This is motivated by an interest in determining how the poles of the impedance function vary with angle of incidence and polarisation.

In all tests, frequency scaling is assumed such that after initial definition of the frequencies all frequencies are scaled by the scaling factor ω_{sc}, such that $\omega = \omega'/\omega_s$ where ω' is a data entry frequency. Frequency scaling applies to the sampling frequencies as well as the initial estimates of the poles and residues (including the e term, if present, which scales as $1/frequency$).

There are some other matters that require definition. The first is concerned with the presence or absence of the d and e terms. In general, for impedance functions of unknown type or order these should both be present and this will be considered the default option. The second is concerned with the initial designation of real vs imaginary poles defined by the variables P and R (as defined in section 2.5) prior to initiating the iteration process. The default option is to assume P is maximum, i.e. that for N defined even, $R = 0$ and for N defined odd, $R = 1$. This will be assumed unless indicated otherwise. As iteration proceeds, these numbers evolve.

In these examples all weightings will be uniform with $w(\omega_i) = 1$. Finally, there is the matter of initial pole estimates. These can either be defined, under user control, by a specification file or they can be defined by a simple distribution as follows (the default option),

For real poles,

$$q_i = -i\,\frac{(\omega_{min} + \omega_{max})}{2R} \quad \text{, for } i = 1, R \tag{3-1}$$

and for complex poles,

$$q_i = -0.001\,\omega_{max} \pm j\left(\omega_{min} + i\,\frac{(\omega_{max} - \omega_{min})}{P+1}\right) \quad \text{, for } i = 1, P \tag{3-2}$$

where ω_{min} and ω_{max} are the smallest and largest frequencies in the sampling range. Note that these are also scaled by the scaling factor ω_{sc}.

3.2 Example 1. A simple 5-pole response in the absence of noise

In this example, two prototype impedance functions are considered, both of which have a zero at the origin. The first has two poles,

$$Z_{3,2}(s;\xi_1,\xi_2,\xi_3,\delta_1) = \frac{s(1+\xi_1 s+\xi_2 s^2)}{1+(\xi_1+\delta_1)s+(\xi_2+\xi_3)s^2} \tag{3-3}$$

with $s = j\omega$ and positive real coefficients ξ_1, ξ_2, δ_1 and ξ_3. The function is positive real if $\xi_1\xi_3 - \delta_1\xi_2 \geq 0$ and $\delta_1^2\xi_2 - \delta_1\xi_1\xi_3 + \xi_3^2 \geq 0$.

The second a single real pole,

$$Z_{1,1}(s;R_1,L_1) = \frac{R_1 s}{s+R_1/L_1} \tag{3-4}$$

A sum of two terms is now constructed, one of type $Z_{3,2}(s)$ and one of type $Z_{1,1}(s)$,

$$Z_{in}(s) = Z_{3,2}(s;\xi_{11},\xi_{21},\xi_{31},\delta_{11}) + Z_{3,2}(s;\xi_{12},\xi_{22},\xi_{32},\delta_{12}) + Z_{1,1}(s;R_1,L_1) \tag{3-5}$$

with parameters,

$\xi_{11} = 0.1$	$\xi_{21} = 0.1$	$\xi_{31} = 1.0$	$\delta_{11} = 0.01$
$\xi_{12} = 0.05$	$\xi_{22} = 0.05$	$\xi_{32} = 0.2$	$\delta_{12} = 0.2$
$L_1 = 2.0$	$R_1 = 1.0$		

This gives rise to the complex poles $a_i = -0.5 \pm j\,1.9364917$, $-0.05 \pm j\,0.95215068$ and the real pole at $a_i = -0.5$.

Here the number of samples $N_f = 50$ with $\omega_i = 0.01 + (i-1)\,0.04$ for $i = 1, N_f$. It is important to define the sampling range in any tests since predictions of pole estimates can be strongly influenced by this. There is no frequency scaling here; i.e. the scaling factor $\omega_{sc} = 1$. It is also assumed that the model order is known and defined by $N = 5$ with both d and e terms present and required.

Gustavsen's first variant (section 2.7)

Using this variant no iteration is required and the correct poles are determined on first use to 8 significant figures and a relative error $E_2^{(rel)} = 1.0 \times 10^{-12}$.

Gustavsen's variant 2a (section 2.8)

Using this variant no iteration is required and the correct poles are determined on first use to 8 significant figures and a relative error $E_2^{(rel)} = 9.5 \times 10^{-11}$.

Single null eigenvalue method (section 2.5)

Using this variant no iteration is required and the correct poles are determined on first use to 8 significant figures and a relative error $E_2^{(rel)} = 3.4 \times 10^{-11}$.

Modified second variant (section 2.7)
Using this variant no iteration is required and the correct poles are determined on first use to 8 significant figures and a relative error $E_2^{(rel)} = 1.0 \times 10^{-12}$. Result is identical to Gustavsen's first variant since $\tilde{d}$ is set to unity with this data.

Null-space method (section 2.6)
This method is employed with $L = 1000$ for null-eigenvector selection. Using this variant no iteration is required and the correct poles are determined on first use to 8 significant figures and a relative error $E_2^{(rel)} = 3.4 \times 10^{-11}$. Result is identical to the single null eigenvalue method since only one near-null eigenvector is detected.

In this example there is good accuracy in all variants. Figure 3-1 shows the real and imaginary parts of the specified impedance function input data and the reconstructed impedance using the predicted poles and residues. This is taken from Gustavsen's first variant (the other variants produce the same prediction).

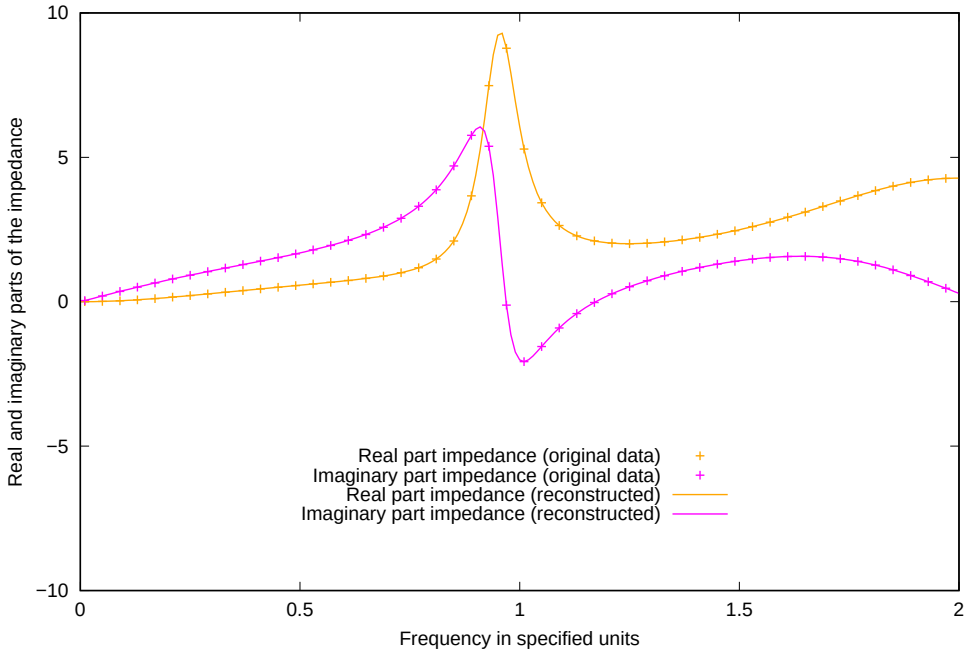

Figure 3-1: *Impedance function for example 1 (results taken from Gustavsen's first variant)*

3.3 Example 2. Single real pole in the presence of noise

This example comes from [2]. Unfortunately no sampling information or details of the noise distribution were provided there so additional assumptions have to be made. The test impedance function is defined by,

$$Z_{in}(s_i) = \frac{-\alpha}{s_i - \alpha} N_i \tag{3-6}$$

for a constant $\alpha = -2\pi f_\alpha$ where $f_\alpha = 100$ kHz. I.e. the pole is located at $a_1 = -6.28318 \times 10^5$. The factor N_i represents multiplicative uncorrelated noise defined by,

$$N_i = 1.0 + 0.02(n_i - 1/2) \tag{3-7}$$

where n_i is uniformly distributed uncorrelated white noise with values over the interval $[0, 1]$ from a standard library routine. N_i provides a $\pm 1\%$ random deviation to the impedance function as described in [2].

Sampling assumes a logarithmic distribution of frequencies with frequencies defined by,

$$\omega_i = \omega_{min} \times 10^{(i-1)\delta} \quad \text{for } i = 1 \text{ to } N_f \tag{3-8}$$

where $N_f = 100$, $\delta = (\log_{10}(\omega_{max}/\omega_{min}))/(N_f - 1)$, $\omega_{min} = 2\pi f_{min}$ and $\omega_{max} = 2\pi f_{max}$. For the example $f_{min} = 10$ Hz and $f_{max} = 1$ MHz.

The scaling frequency $\omega_{sc} = 2\pi \times 10^5$ and an initial pole estimate $q_1 = 2\pi \times 10$ as in [2]. The model order is assumed known so take $N = 1$ with no d and e terms.

Gustavsen's first variant (section 2.7)

Using this variant the algorithm converges to a pole estimate -6.2860×10^5 (100.0 kHz) after 4 iterations. On the first iteration the pole is estimated at -4.181×10^3 (665 Hz). The relative error after 5 iterations has a value $E_2^{(rel)} = 5.931 \times 10^{-3}$. This is also the value for $E_1^{(rel)}$.

Gustavsen's variant 2a (section 2.8)

Convergence is faster, as expected from [2]. On the first iteration the pole is estimated at -4.353×10^5 (69.3 kHz) with a value of $\tilde{d} = 0.00446$. On the second iteration the pole is estimated at -6.288×10^5. On the third iteration convergence is achieved with a pole estimate -6.2891×10^5 (100.1 kHz) and a value $E_2^{(rel)} = 5.929 \times 10^{-3}$.

Single null eigenvalue method (section 2.5)

Using this variant the algorithm converges to a pole estimate -6.2881×10^5 (100.1 kHz) after 4 iterations with a value for $E_2^{(rel)} = 5.930 \times 10^{-3}$. On the first iteration the pole is estimated at -4.237×10^3 (674 Hz). This is interesting in that $\tilde{d}$ is calculated on this iteration at $\tilde{d} \approx 0.0149$, not as small as in Gustavsen's second variant but still significantly less than one. Despite this, the first iteration pole estimate is similar to that in Gustavsen's first variant.

Modified second variant (section 2.7)

Using this variant the algorithm converges to a pole estimate -6.2860×10^5 (100.0 kHz) after 4 iterations with a value for $E_2^{(rel)} = 5.931 \times 10^{-3}$. On the first iteration the pole is estimated at -4.181×10^3 (665 Hz) with a value of $\tilde{d} \approx 0.0149$ (expected, since the value comes from the null eigenvalue result).

Null-space method (section 2.6)

This method is employed with $L = 1000$ for null-eigenvector selection. Only one null eigenvector is selected by the algorithm so the results default to those of the single null eigenvalue method as described above.

In this example there is also good accuracy in all variants. Figure 3-2 shows the real and imaginary parts of the specified impedance function input data and the reconstructed impedance using the predicted poles and residues. Here the frequency in Hz is plotted on a log scale. This is taken from Gustavsen's first variant (the other variants produce the same prediction).

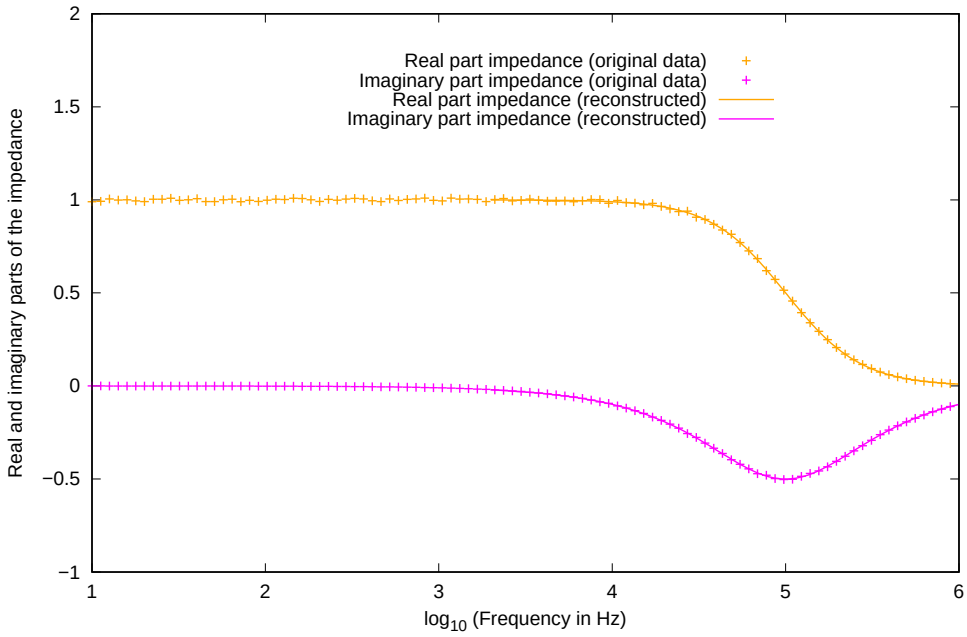

Figure 3-2: *Impedance function for example 2 (results taken from Gustavsen's first variant)*

3.4 Example 3a. Gustavsen 18 pole example

This example attempts to duplicate the impedance function specification of example 4 of [1] (though sampling details were unspecified there). Starting pole positions are defined using (3-2). The model order is assumed known with $N = 18$ with d and e terms required. The impedance function takes the form,

$$Z_{in}(s) = \sum_{i=1}^{N} \frac{r_i}{s - a_i} + d + e\,s \tag{3-9}$$

where the poles and residues are tabulated in Hz:

a_i (Hz)	r_i (Hz)
-4500	-3000
-41000	-83000
-100.0 ±j 5000.0	-5.0 ±j 7000.0
-120.0 ±j 15000.0	-20.0 ±j 18000.0
-3000.0 ±j 35000.0	6000.0 ±j 45000.0
-200.0 ±j 45000.0	40.0 ±j 60000.0
-1500.0 ±j 45000.0	90.0 ±j 10000.0
-500.0 ±j 70000.0	50000.0 ±j 80000.0
-1000.0 ±j 73000.0	1000.0 ±j 45000.0
-2000.0 ±j 90000.0	-5000.0 ±j 92000.0
d (dimensionless)	e (Hz^{-1})
0.2	2.0×10^{-5}

Uniform sampling is assumed with $N_f = 200$ samples over the range $f_{min} = 500$ Hz to $f_{max} = 100$ kHz. Frequency scaling is defined with $\omega_{sc} = 2\pi f_{max}$.

Note that the response function defined with these residues, taken from [1], is not passive. The residues, as defined, give rise to an impedance function with real part that takes positive and negative values and an imaginary part that remains positive. Such a function is by definition not positive real (PR). However this still provides a valid test since (for better or worse) there is no part of the algorithm in any variant of the method that imposes the PR condition.

Gustavsen's first variant (section 2.7)

Using this variant the algorithm arrives at pole positions and residues with high accuracy on the first iteration. At this point the complex poles are correctly specified to 8 significant figure accuracy, the real poles to 7 significant figure accuracy and the relative error, $E_2^{(rel)} = 1.5 \times 10^{-7}$. After a few more iterations convergence is achieved with a relative error $E_2^{(rel)} \sim 10^{-10}$ and poles and residues are all correctly reproduced to 8 significant figure accuracy. Figure 3-4 shows both $log_{10}(E_2^{(rel)})$ and $log_{10}(E_1^{(rel)})$ plotted as a function of the iteration number. Figure 3-3 illustrates the sampled and reconstructed real and imaginary parts of

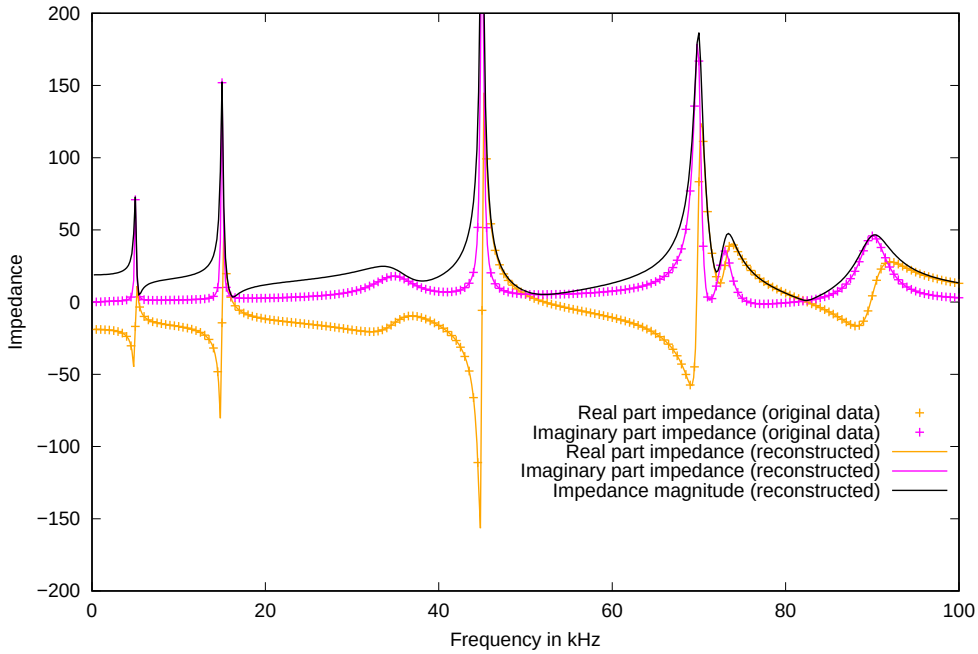

Figure 3-3: *Impedance function for example 3a. Real and imaginary part of impedance and its magnitude as a function of frequency in kHz.*

the impedance function. The magnitude of the reconstructed impedance function is also shown.

Gustavsen's variant 2a (section 2.8)

Using this variant the algorithm converges more slowly. It takes about 4-5 iterations to obtain a relative error $E_2^{(rel)} \sim 10^{-7}$ and no uniform convergence occurs after this point. No value for $E_2^{(rel)}$ is obtained below $\sim 10^{-8}$. This is associated with 7 significant figure accuracy to the residues. The pole positions are predicted with higher accuracy to about 7 or 8 significant figures, as before. Figure 3-5 shows $log_{10}(E_2^{(rel)})$ and $log_{10}(E_1^{(rel)})$ plotted as a function of the iteration number.

Single null eigenvalue method (section 2.5)

Using this variant the algorithm takes 3 iterations to obtain a relative error $E_2^{(rel)} = 5.5 \times 10^{-6}$ after which there is no uniform convergence. No value for $E_2^{(rel)}$ is obtained significantly below $\sim 10^{-6}$ over 30 iterations. Complex poles are accurate to 7 significant figure. Real poles are accurate to 5 significant figures. Residues are accurate to about 7 significant figures for the complex poles and 4 significant figures for the real poles.

Modified second variant (section 2.7)

This method closely tracks the accuracy of Gustavsen's first variant. After the first iteration the relative error $E_2^{(rel)} = 1.5 \times 10^{-7}$, but after the second $E_2^{(rel)} = 5.0 \times 10^{-10}$. The error oscillates with values of the order of $\sim 10^{-10}$ thereafter.

Null-space method (section 2.6)
This method tracks the single null eigenvalue method for accuracy. After 4 iterations the relative error $E_2^{(rel)} \sim 10^{-6}$ Figure 3-6 shows $log_{10}(E_2^{(rel)})$ and $log_{10}(E_1^{(rel)})$ plotted as a function of the iteration number. Only one null eigenvector is selected at each algorithm stage so differences in the precise values of $E_2^{(rel)}$ and pole and residue predictions is attributed to machine round-off errors.

3.5 Example 3b. Gustavsen 18 pole example with PR residues

This example employs the same poles as above, taken from example 4 of [1], but the residues are redefined to ensure a passive response function whose magnitude is similar to the one presented there. There are no simple conditions on residues to ensure a function is PR, but it is easier to construct one if the real parts are all positive, which is the case here. Starting pole positions are defined using (3-2). The model order is assumed known; i.e. $N = 18$ with d and e terms required and the impedance function takes the same form. The poles and new residues, tabulated in Hz, are given in the table:

a_i (Hz)	r_i (Hz)
-4500	45000
-41000	83000
-100.0 ±j 5000.0	2500.0 ∓j 3000.0
-120.0 ±j 15000.0	10000.0 ∓j 2000.0
-3000.0 ±j 35000.0	65000.0 ±j 6000.0
-200.0 ±j 45000.0	25000.0 ±j 40.0
-1500.0 ±j 45000.0	10000.0 ±j 90.0
-500.0 ±j 70000.0	80000.0 ∓j 5000.0
-1000.0 ±j 73000.0	45000.0 ±j 1000.0
-2000.0 ±j 90000.0	92000.0 ∓j 5000.0
d (dimensionless)	e (Hz^{-1})
0.2	2.0×10^{-5}

The previous uniform sampling is assumed with $N_f = 200$ over the range $f_{min} = 500$ Hz to $f_{max} = 100$ kHz and $\omega_{sc} = 2\pi f_{max}$.

Using Gustavsen's first variant (section 2.7) the algorithm arrives at pole positions and residues with high accuracy on the first iteration. At this point the complex poles are correctly specified to 8 significant figure accuracy, the real poles to 6 significant figure accuracy and the relative error $E_2^{(rel)} = 3.6 \times 10^{-7}$. After a few more iterations convergence is achieved with a relative error $E_2^{(rel)} \sim 10^{-10}$ and the predicted poles and residues are accurate to 8 significant figures. Figure 3-7 illustrates the sampled and reconstructed real and imaginary parts of the impedance function. The magnitude of the reconstructed impedance function is also shown. Note that the real part is always positive here. The other variants behave much the same as described in the previous example.

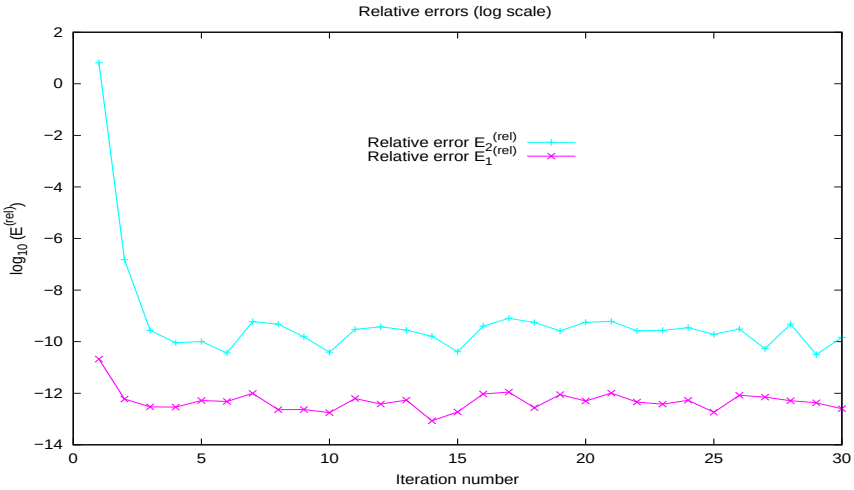

Figure 3-4: *Example 3a, Gustavsen's first variant.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration count*

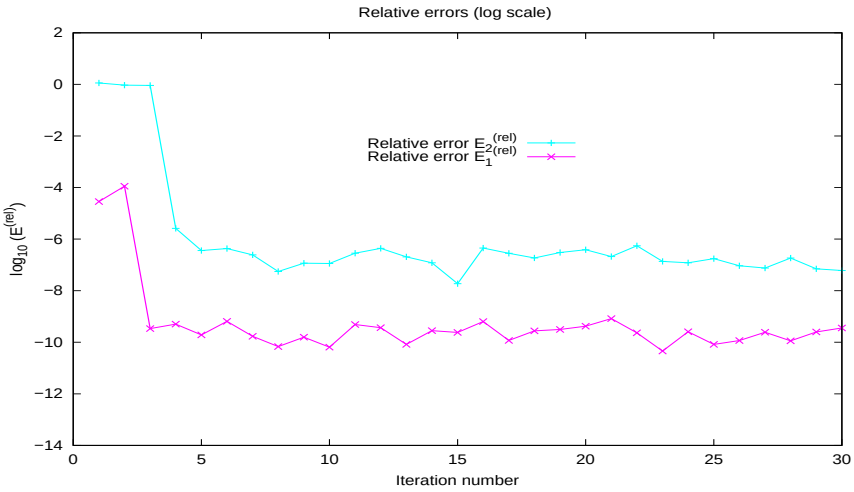

Figure 3-5: *Example 3a, Gustavsen's variant 2a.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration count*

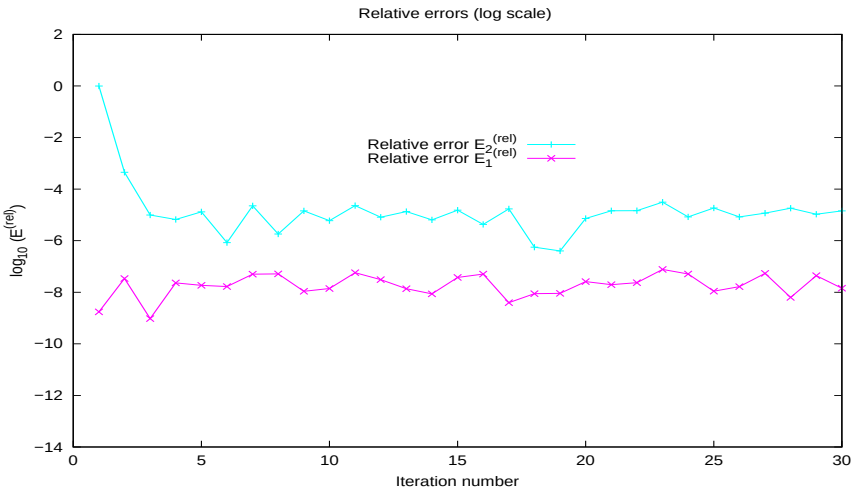

Figure 3-6: *Example 3a, Null space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration count*

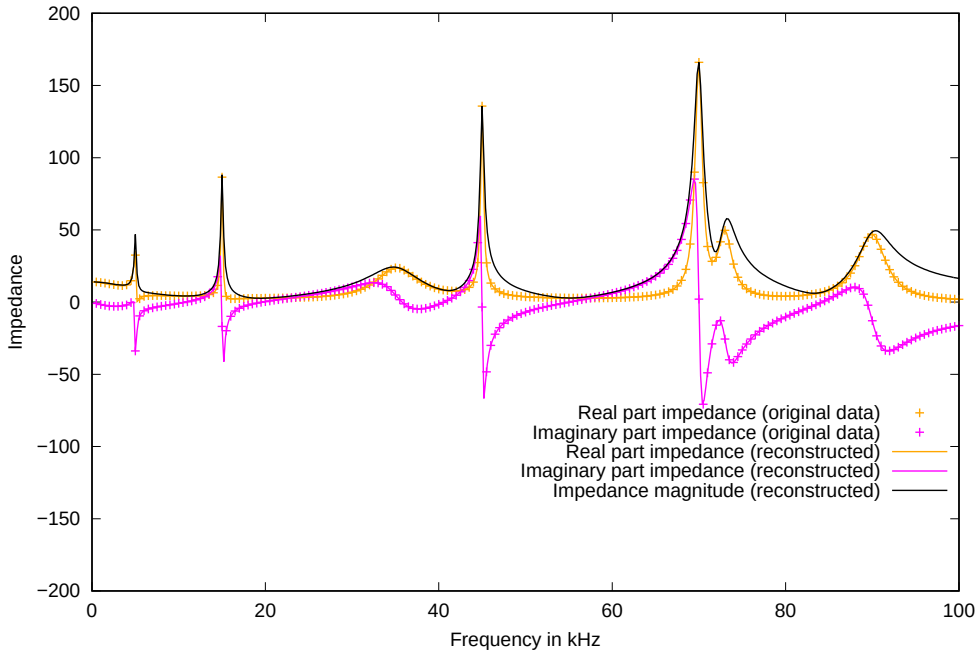

Figure 3-7: *Impedance function for example 3b. Real and imaginary part of impedance and its magnitude as a function of frequency in kHz.*

3.6 Example 4. Gustavsen's 18 pole example with model order overestimated

This example is specified as in example 3a, above, except that the model order is taken to be $N = 20$ or $N = 40$ (d and e assumed required). This test was also made in [1] and results in non-uniqueness of the poles and residues and a matrix Q, in Gustavsen's first variant and a matrix Ω in the second variant, that do not possess inverses. This is where the Moore-Penrose definitions are required, using the SVD.

Gustavsen's first variant (section 2.7)

When $N = 20$, convergence is slow but fairly uniform until the relative error $E_2^{(rel)} \sim \times 10^{-8}$. Approximately 35 iterations were required. See figure 3-8. The dimension of the matrix Q is $2N+2 = 42$ but its rank is 40 so the Moore-Penrose generalised inverse contains two zeros in the diagonal matrix Σ^{-1}. The algorithm predicts two pole pairs at $-500.0098 \pm j70000.000$ and $-499.9902 \pm j70000.000$ and divides the residue equally between them with respective residues $25000.058 \pm j40000.015$ and $24999.942 \pm j39999.985$. This phenomenon was not reported in [1] where the mechanism there was the allocation of extraneous poles with near zero residues. The other poles and residues are accurate to 8 significant figures. The fitting of the data to the predictions is illustrated in figure 3-13, which may be considered good.

When $N = 40$, convergence is poor with a minimum relative error $E_2^{(rel)} \sim 5.3 \times 10^{-2}$ which occurs at around 25 iterations. The error subsequently degrades and becomes erratic for further iterations. See figure 3-8. At this point the rank of the matrix Q is 60 (dimension

82). The two real poles are correctly estimated to 6 significant figures but their residues are poorly estimated (accuracy about 90%). The complex poles are split, with multiple predicted poles bunched together. For example, the initial-data pole pairs at $-2000 \pm j90000$ are represented by three pole pairs at $-1998.4378 \pm j89993.860$, $-1995.3343 \pm j90004.321$ and $-2006.0563 \pm j90001.795$. No spurious poles with near zero residues were generated. The poor fitting of the data to the predictions is illustrated in figure 3-14.

Gustavsen's variant 2a (section 2.8)

When $N = 20$, convergence is very poor and erratic with a minimum relative error $E_2^{(rel)} \sim 10^{-1}$. See figure 3-9. At about 20 iterations $\tilde{d} \approx 0.979$ (it should be 1). The rank of the matrix Ω is 40 and the algorithm correctly locates the complex poles but predicts 4 real poles that are spread around the correct values and the residues are very poorly estimated.

When $N = 40$, convergence does not occur and $E_2^{(rel)} \sim 1$. See figure 3-9. The rank of the matrix Ω varies from around 56 to 60. On the other hand, the relative error $E_1^{(rel)}$ is quite small at around 10^{-6}. The error E_1 (or more precisely $E_1^{(rms)}$) is the one illustrated in [2]. Figure 3-15 shows the poor fitting of the data to the predictions in this case. This very clearly demonstrates that E_1 does not provide a good indication of accuracy when the model order is overestimated.

Gustavsen's variant 2b (section 2.8)

When $N = 20$, convergence is slow but fairly uniform until the relative error $E_2^{(rel)} \sim \times 10^{-7}$. See figure 3-10. The Accuracy and convergence is similar to Gutavsen's first variant. The parameter $\tilde{d}$ converges to 1 uniformly and to about 7 significant figures after iteration 30.

When $N = 40$ convergence does not occur. See figure 3-10. The parameter $\tilde{d}$ does not converge to 1 and $E_2^{(rel)}$ remains ~ 1. Figure 3-16 shows the poor fitting of the data to the predictions in this case.

Single null eigenvalue method (section 2.5)

When $N = 20$, convergence is poor and erratic with a minimum relative error $E_2^{(rel)} \sim 2.6 \times 10^{-4}$ at iteration 8, but $E_2^{(rel)} \sim 1$ in some later iterations. This is symptomatic of taking the wrong 'direction' from within the null-space. At iteration 8, the complex poles are predicted accurate to about 7 significant figures, but the real poles only to about 3 significant figures. In this example two spurious real poles are predicted with values $\sim 10^6$ and large residues of opposite signs. The complex residues are accurate to about 3 significant figures.

When $N = 40$, there is no convergence with a relative error that usually takes a value $E_2^{(rel)} \sim 1$ with occasionally smaller values when, presumably, a correct direction is chosen in the null space. For example at iteration 40 a value $E_2^{(rel)} \approx 0.088$ was found. At this point 14 real poles and 13 complex pole pairs were predicted. This contained a mixture of spurious and bunched poles. This method is only intended for demonstration purposes and results are not plotted.

Modified second variant (section 2.7)
Results here were very similar to Gustavsen's first variant when $N = 20$. However, when $N = 40$ the relative error never gets smaller than $E_2^{(rel)} \sim 10^{-1}$. See figure 3-11. Here $\tilde{d}$ is frequently calculated, via the null-eigenvalue method, to be out of bounds. Resetting $\tilde{d} \to 1$, as defined in the algorithm, in these circumstances results in the discontinuous nature of the plots.

Null-space method (section 2.6)
When $N = 20$, three eigenvectors are automatically selected to form the null-space. After about 5 iterations the relative error $E_2^{(rel)} \sim 10^{-5}$ after which there is variation between about 10^{-5} and 10^{-6}. See figure 3-12. A spurious pole pair was predicted with value $-1471.2 \pm j51944.0$ with residues two orders of magnitude down on the smallest residue associated with the non-spurious poles. Accuracy of the non-spurious poles is better than about 3 significant figures for all these poles. Figure 3-17 shows the quality of the data fit for the $N = 20$ case.

When $N = 40$, 23 eigenvectors are automatically selected to form the null-space. After just one iteration the relative error $E_2^{(rel)} \approx 4.7 \times 10^{-5}$. There is no uniform convergence and on subsequent iterations $E_2^{(rel)}$ lies between $\sim 10^{-5}$ and $\sim 10^{-6}$. See figure 3-12. Predicted poles are spread with some far from the correct values, with small residues and some close to the correct values with distributed residues. Some of the poles are isolated and predicted correctly to 7 significant figures with residues accurate to about 4 significant figures. Figure 3-18 shows the quality of the data fit for the $N = 40$ case.

When the model order is estimated much higher than the correct order only the null-space method is capable of faithfully representing the impedance function with high accuracy. When the model order is correctly specified Gustavsen's variants are capable of greater accuracy, but the null-space method is still good. It is also important to distinguish between the two error measures E_1, as discussed in [1], and E_2. It is unclear, in [2], which measure is being used. This example demonstrates that E_1 and E_2 have very different properties.

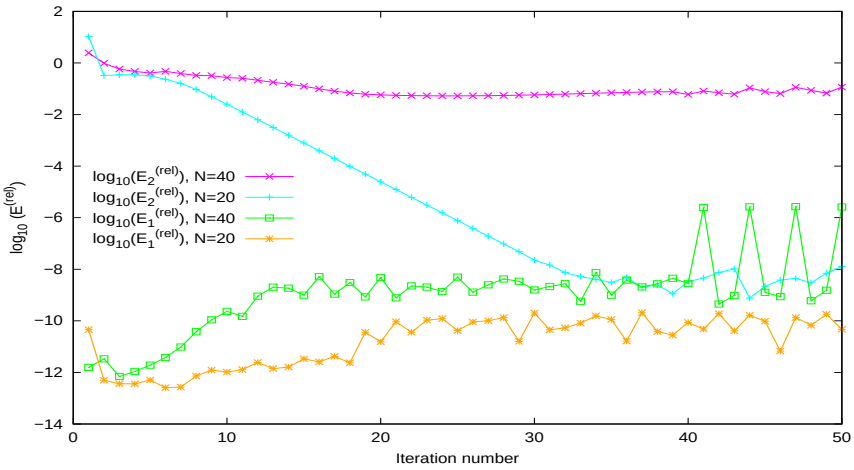

Figure 3-8: *Example 4, Gustavsen's first variant.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

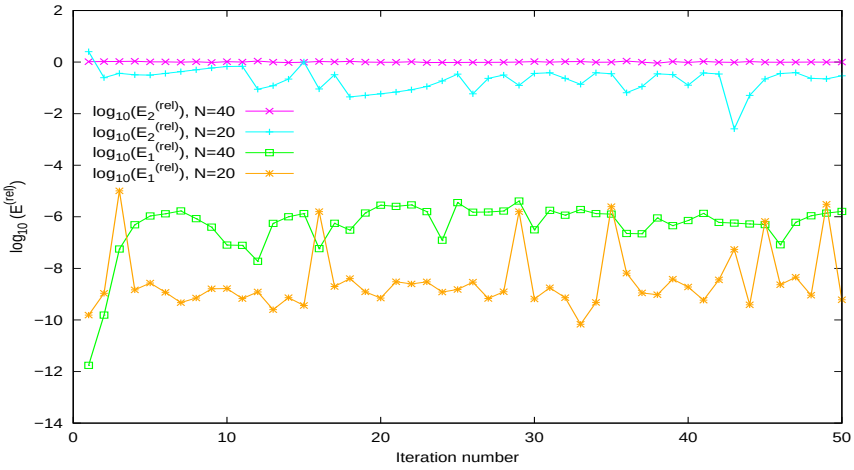

Figure 3-9: *Example 4, Gustavsen's variant 2a.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

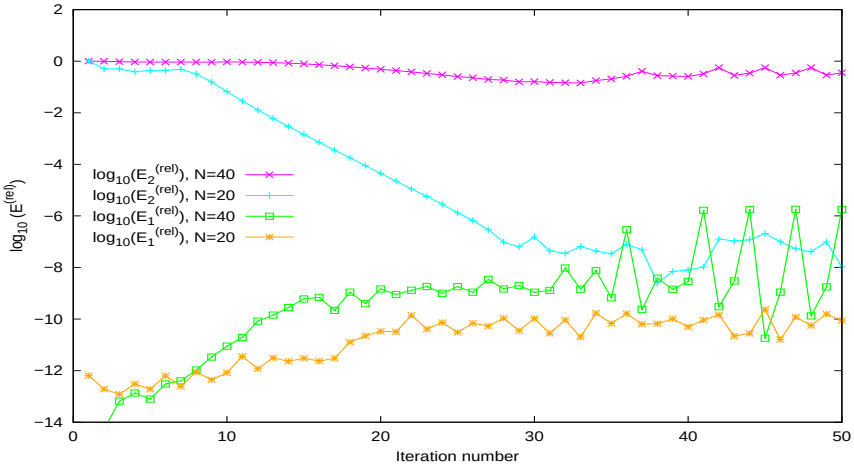

Figure 3-10: *Example 4, Gustavsen's variant 2b.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

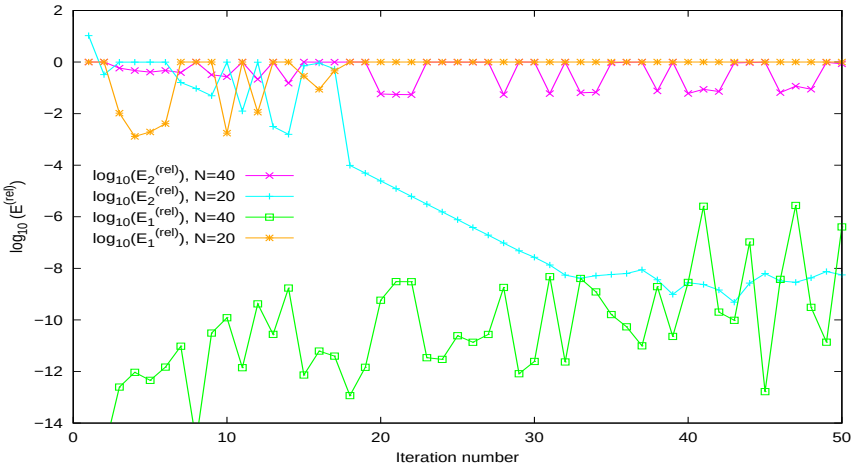

Figure 3-11: *Example 4, Modified second variant.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

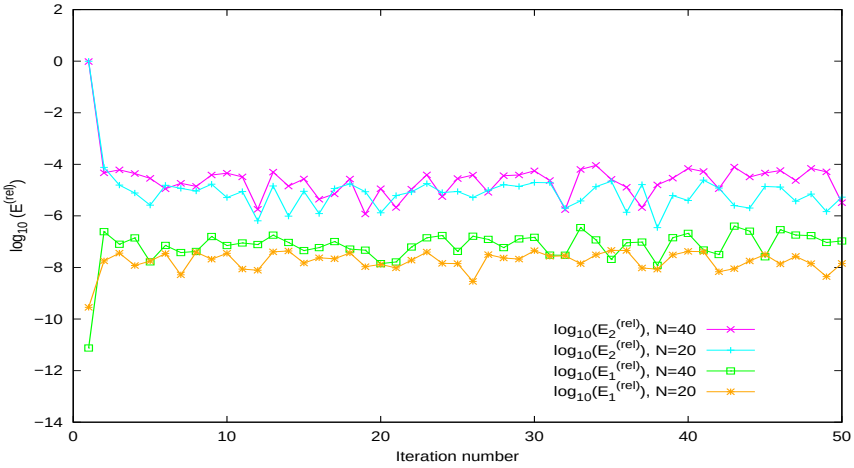

Figure 3-12: *Example 4, Null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

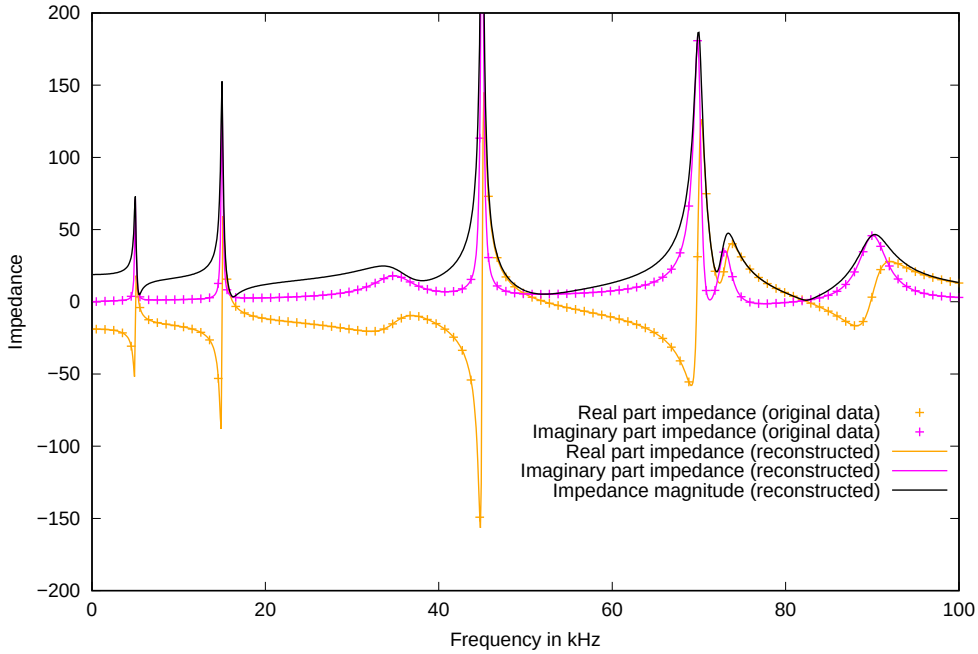

Figure 3-13: *Example 4, Gustavsen's first variant for* $N = 20$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

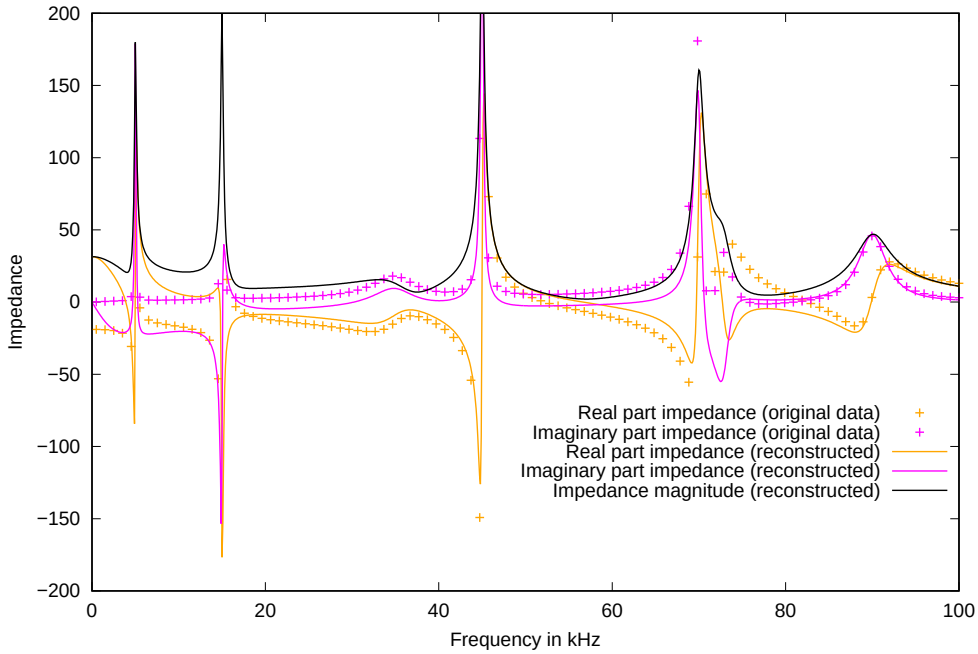

Figure 3-14: *Example 4, Gustavsen's first variant for* $N = 40$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

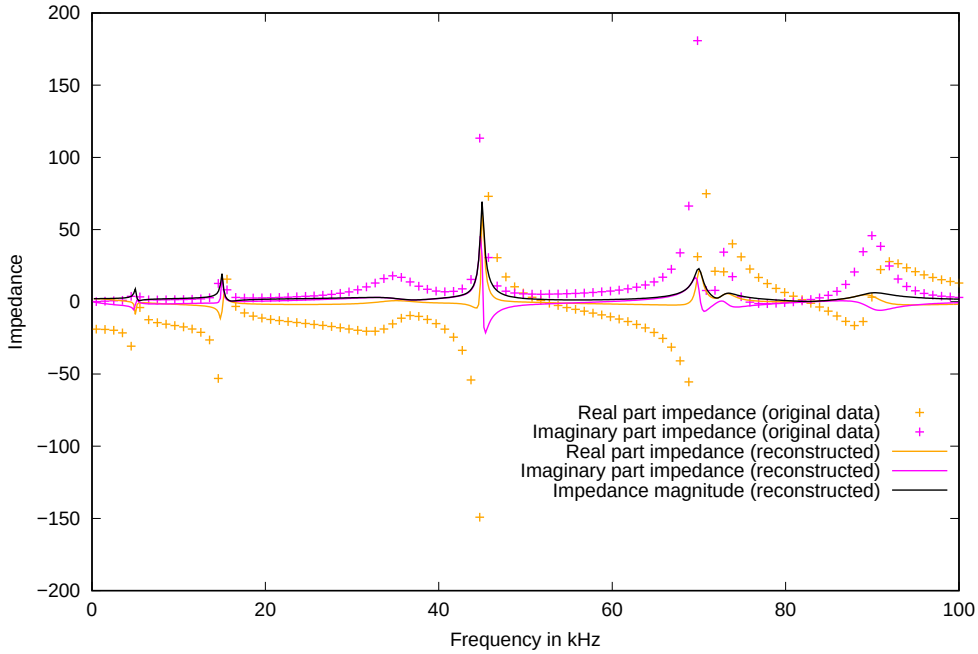

Figure 3-15: *Example 4, Gustavsen's variant 2a for* $N = 40$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

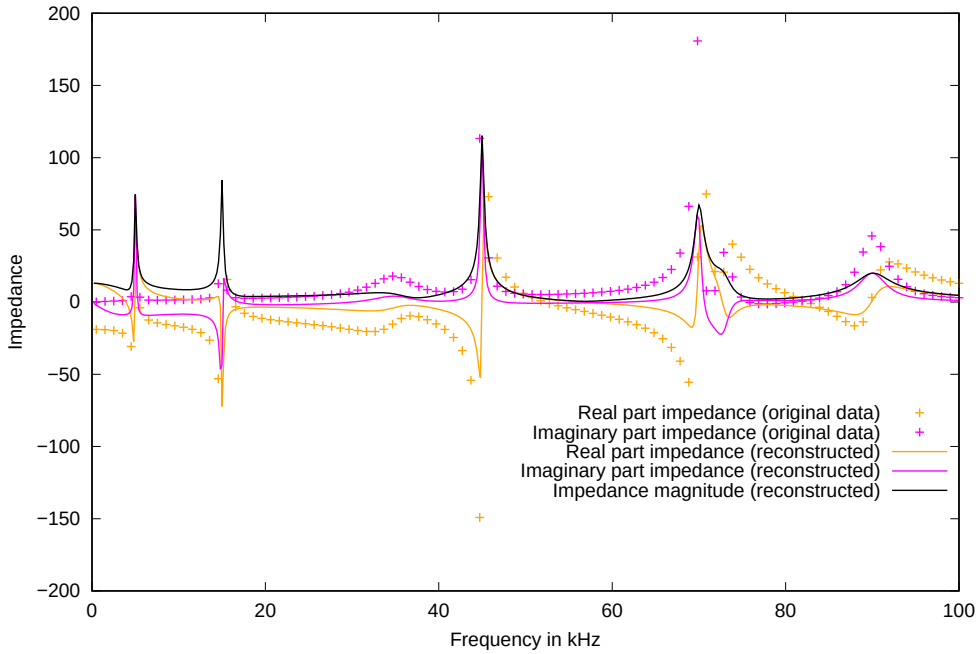

Figure 3-16: *Example 4, Gustavsen's variant 2b for* $N = 40$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

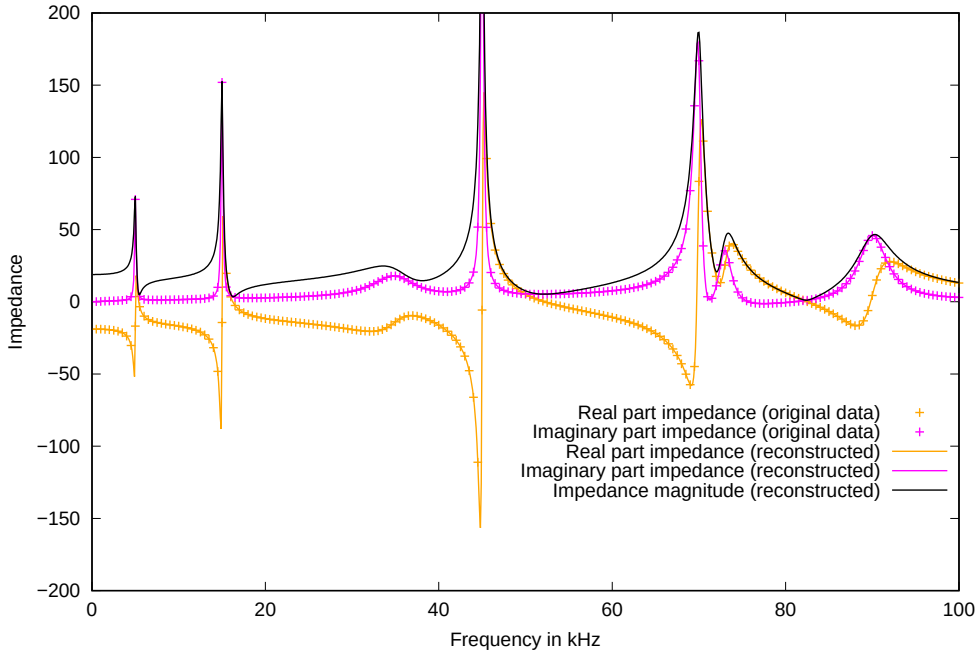

Figure 3-17: *Example 4, null-space method for $N = 20$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

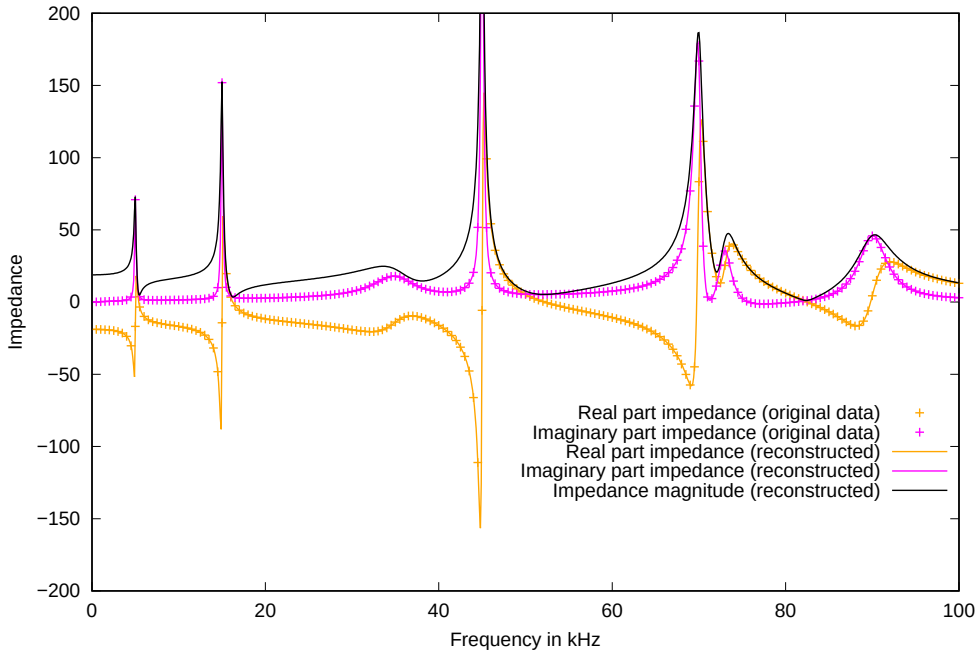

Figure 3-18: *Example 4, null-space method for $N = 40$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

3.7 Example 5. The 18 pole example with model order underestimation

This example is specified as in example 3a, above, except that the model order is taken to be $N = 12$, $N = 14$ or $N = 16$ (d and e assumed required), which is less than defined by the model. This test was performed in [1] for the case $N = 14$ for Gustavsen's first variant. When the model order is taken to be less than defined by a finite order model, the matrices Q (for the first variant) and Ω (in the second variant) usually possess inverses and the Moore-Penrose definitions are generally not required (although this may not be true if the model order is sufficiently large or the real parts of some of the poles are sufficiently negative).

Gustavsen's first variant (section 2.7)

In all three cases, $N = 12$, $N = 14$ or $N = 16$, convergence occurs within 10 iterations with $E_2^{(rel)} \rightarrow E_1^{(rel)}$. This is shown in figure 3-19, which plots $\log_{10}(E_1^{(rel)})$ and $\log_{10}(E_2^{(rel)})$ as a function of iteration number. Plots of the original impedance data and the reconstructed real and imaginary parts of the impedance as a function of frequency are shown in figures 3-25, 3-26 and 3-27. This is consistent with the findings in [1], where good matching to the original data is found provided the number of pole pairs is not less than the number of distinct peaks in the original data.

Gustavsen's variant 2a (section 2.8)

For the cases $N = 12$ and $N = 16$, convergence does not occur. Interestingly, it does occur when $N = 14$. This is shown in figure 3-20. Figures 3-28, 3-29 and 3-30 show the impedance reconstructions after 50 iterations. The convergent case, $N = 14$, provides the best impedance fit and the smallest $E_2^{(rel)}$ and the accuracy of fit is almost the same as Gustavsen's first variant for $N = 14$.

Gustavsen's variant 2b (section 2.8)

Here there is quite rapid convergence for all three cases, $N = 12$, $N = 14$ and $N = 16$. However, $E_1^{(rel)}$ and $E_2^{(rel)}$ remain quite different as shown in figure 3-21. Plots of the original impedance data and the reconstructed real and imaginary parts of the impedance as a function of frequency are shown in figures 3-31, 3-32 and 3-33. Again, the case $N = 14$ provides the best impedance fit and the smallest $E_2^{(rel)}$. Although, for this case $E_1^{(rel)}$ converges to nearly the same value as in variant 2a, the value for $E_2^{(rel)}$ is larger and the impedance fit is not as good.

Null-space method (section 2.6)

For this method, the convergence and errors depend on the threshold value L which controls the size of the null-space. Convergence occurs for all three cases, $N = 12$, $N = 14$ and $N = 16$, but can take a while to achieve. This is illustrated in figures 3-22, 3-23 and 3-24, respectively. These each show predictions made with $L = 10000$, $L = 1000$ (the default value) and $L = 10$. For the cases $N = 14$ and $N = 16$ best results are achieved when $L = 10$ with fit errors, defined by $E_2^{(rel)}$, which are similar to those in Gustavsen's first variant. Figures 3-34, 3-35 and 3-36 show the impedance reconstructions after 50 iterations for the case $L = 10$.

When the model order is underestimated Gustavsen's first variant performs well, the variants 2a and 2b are not robust and the null-space method performs well provided L is relatively small. This suggests that for a model of unknown order some variation of the parameter L should be considered in the null-space method, so as to minimise $E_2^{(rel)}$. If, on the other hand, a reduced order model is specifically required it is probably best to employ Gustavsen's first variant.

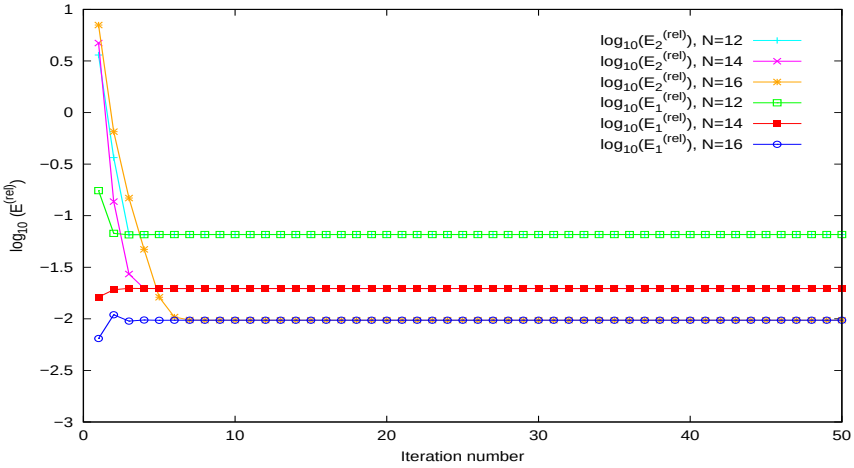

Figure 3-19: *Example 5, Gustavsen's first variant.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 12$, $N = 14$ *and* $N = 16$ *vs iteration count*

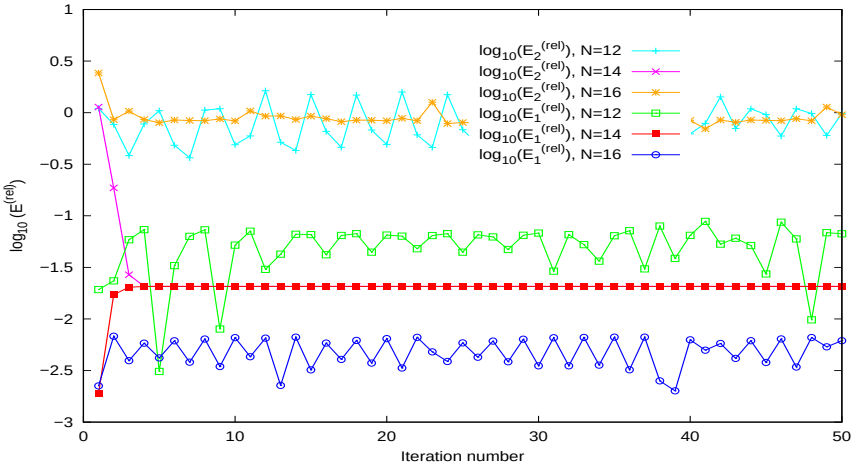

Figure 3-20: *Example 5, Gustavsen's variant 2a.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 12$, $N = 14$ *and* $N = 16$ *vs iteration count*

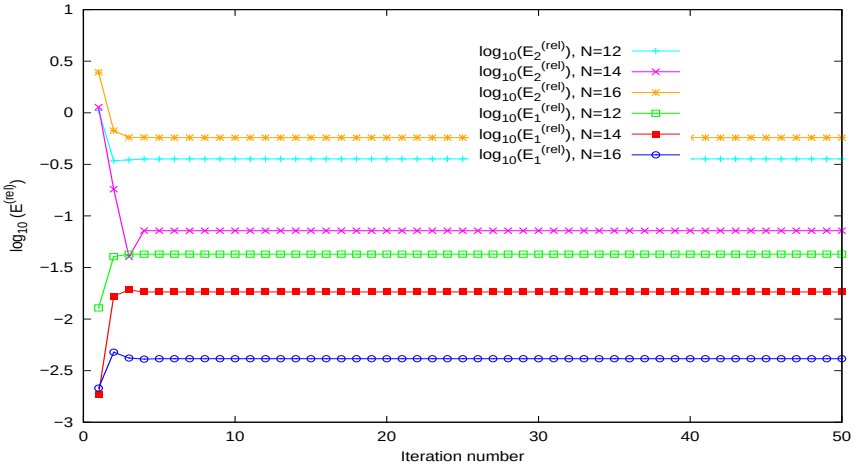

Figure 3-21: *Example 5, Gustavsen's variant 2b.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 12$, $N = 14$ *and* $N = 16$ *vs iteration count*

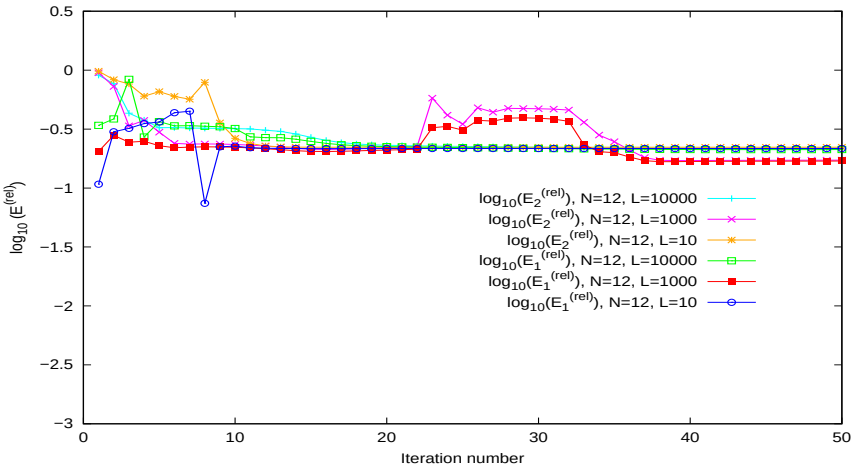

Figure 3-22: *Example 5, Null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 12$ *vs iteration count*

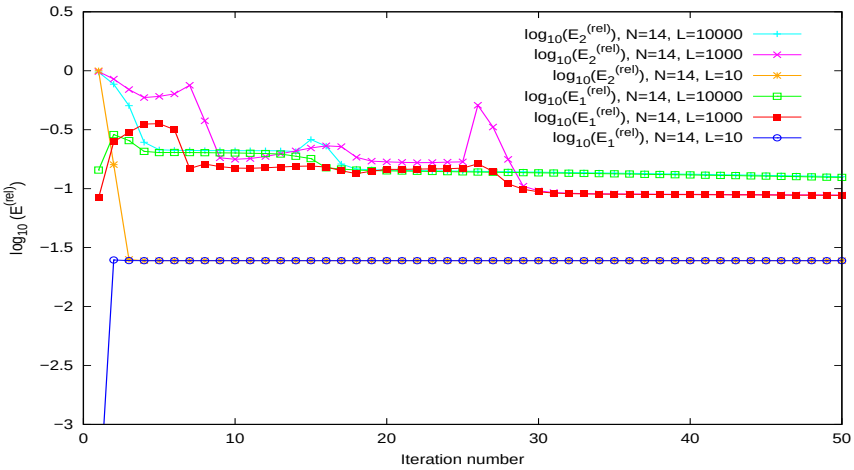

Figure 3-23: *Example 5, Null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 14$ *vs iteration count*

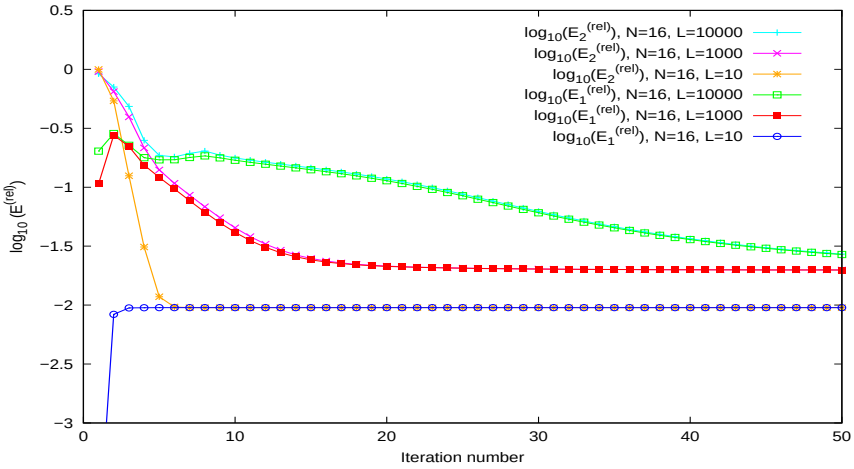

Figure 3-24: *Example 5, Null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 16$ *vs iteration count*

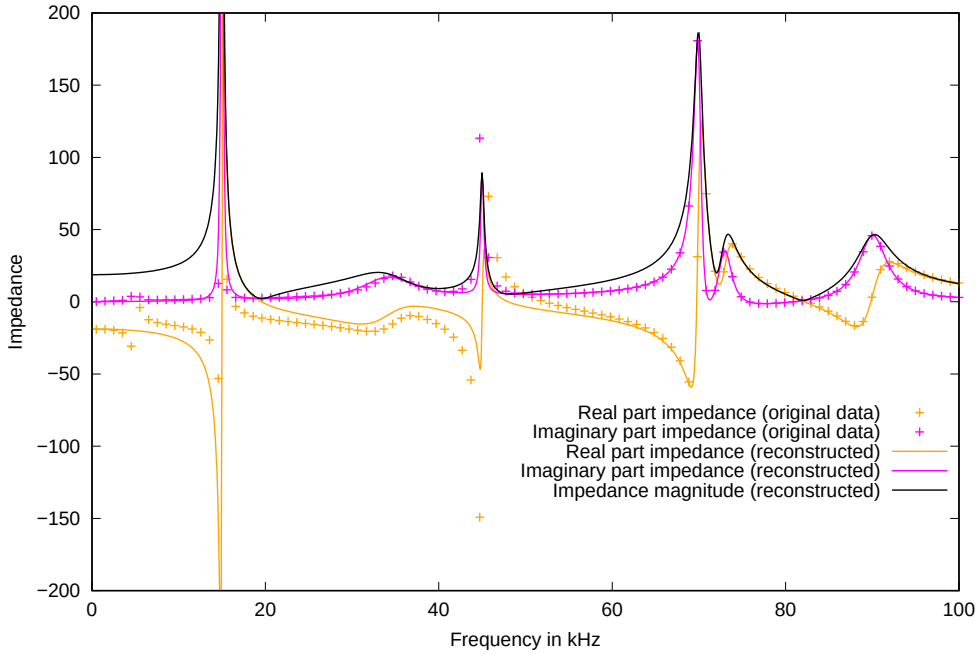

Figure 3-25: *Example 5, Gustavsen's first variant for* $N = 12$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

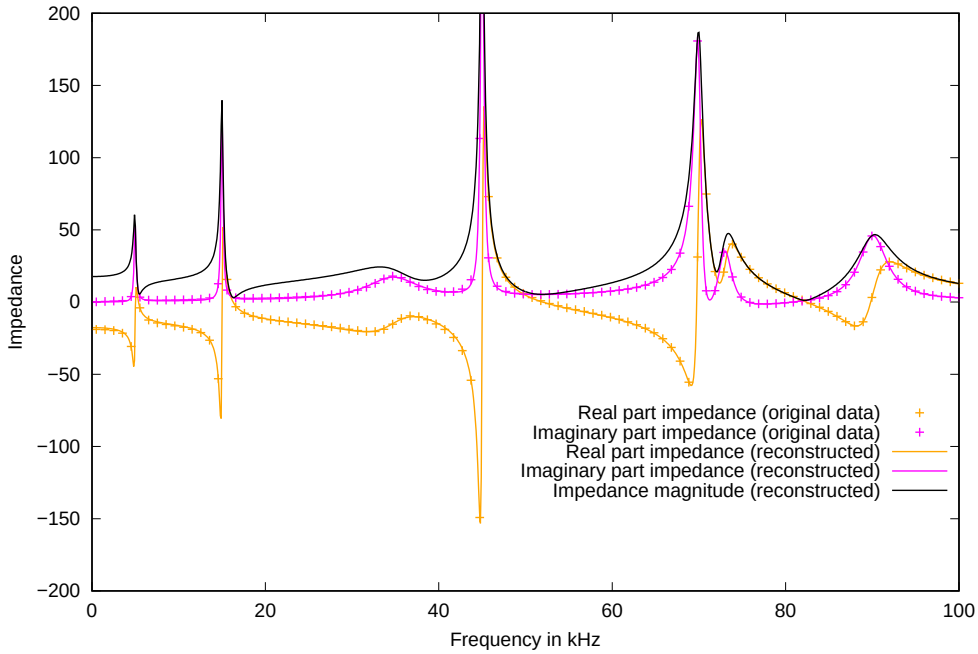

Figure 3-26: *Example 5, Gustavsen's first variant for* $N = 14$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

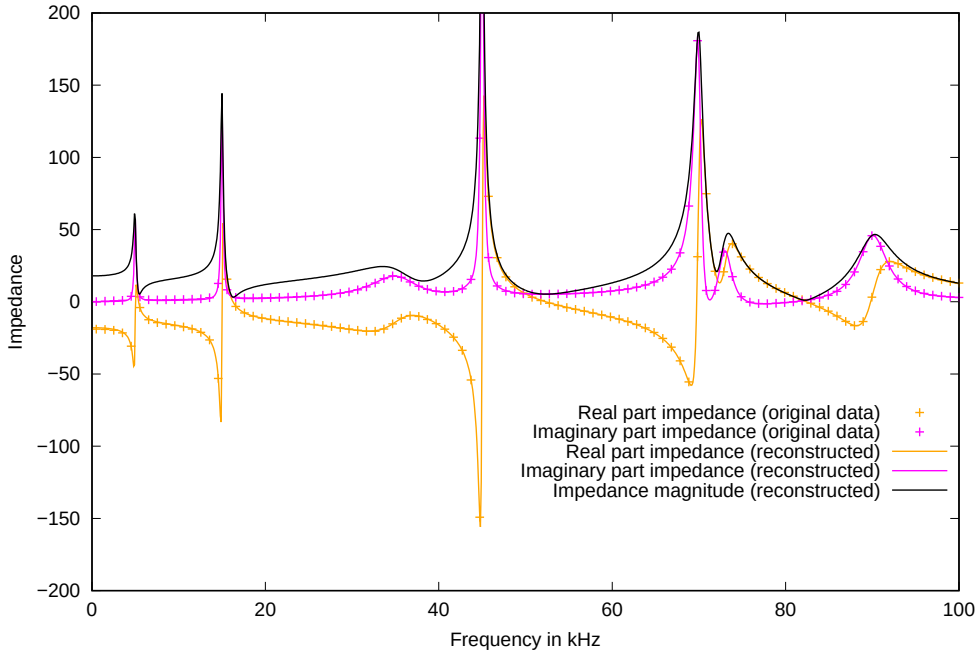

Figure 3-27: *Example 5, Gustavsen's first variant for* $N = 16$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

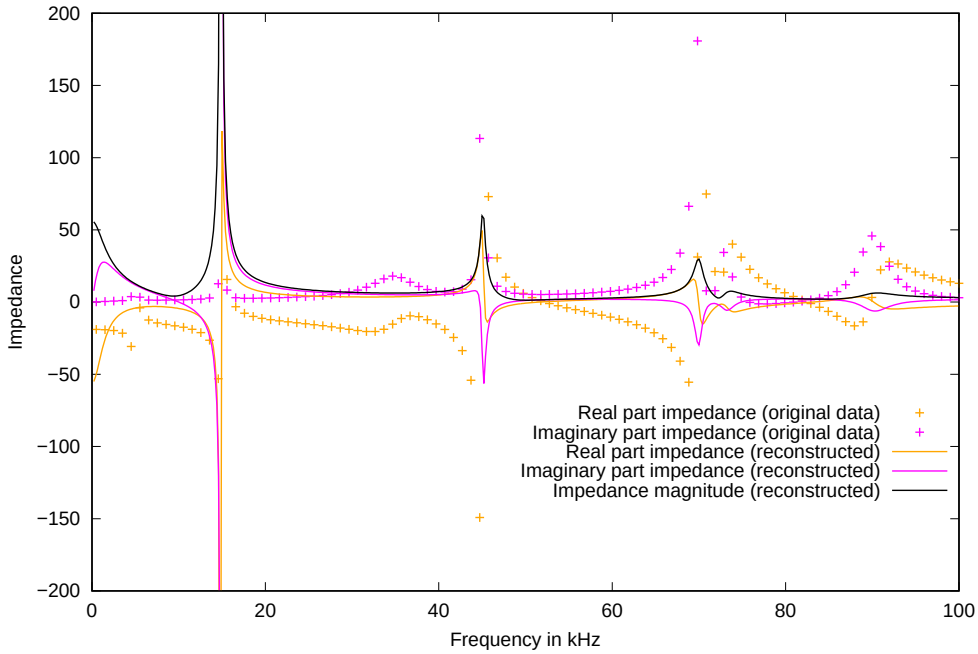

Figure 3-28: *Example 5, Gustavsen's variant 2a for* $N = 12$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

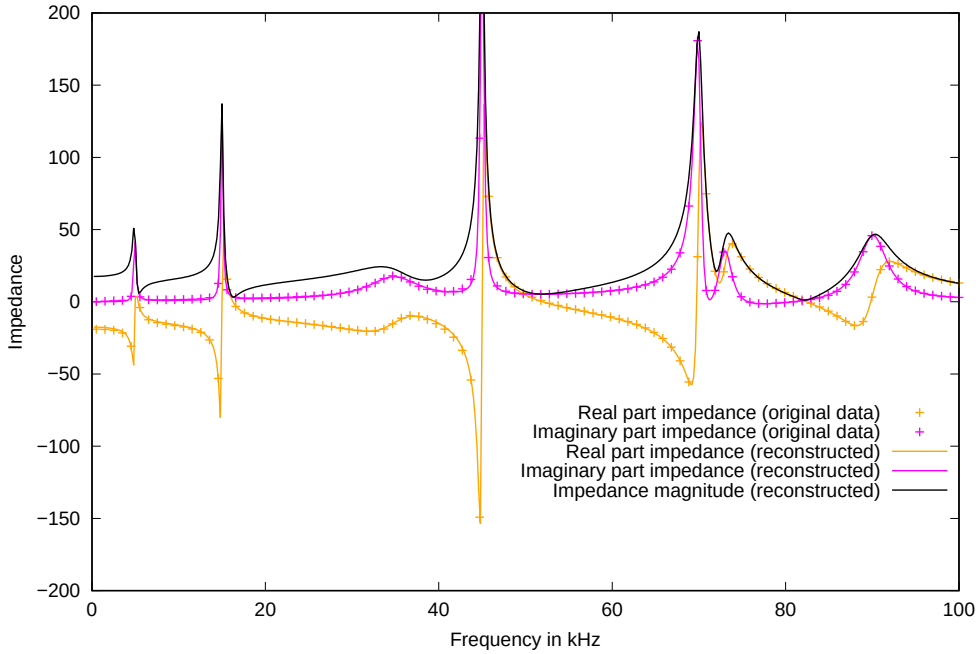

Figure 3-29: *Example 5, Gustavsen's variant 2a for $N = 14$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

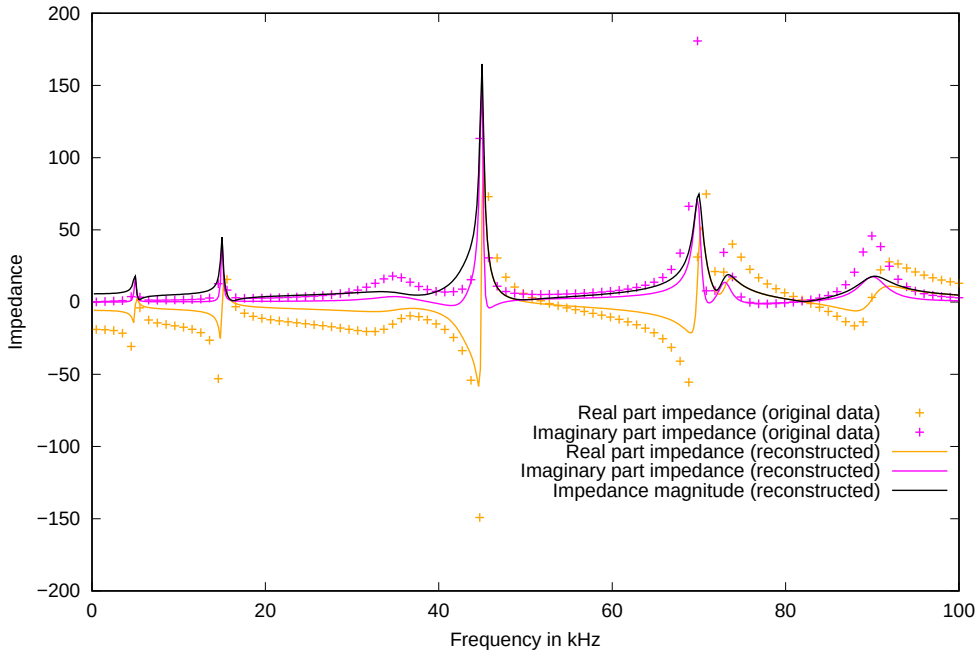

Figure 3-30: *Example 5, Gustavsen's variant 2a for $N = 16$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

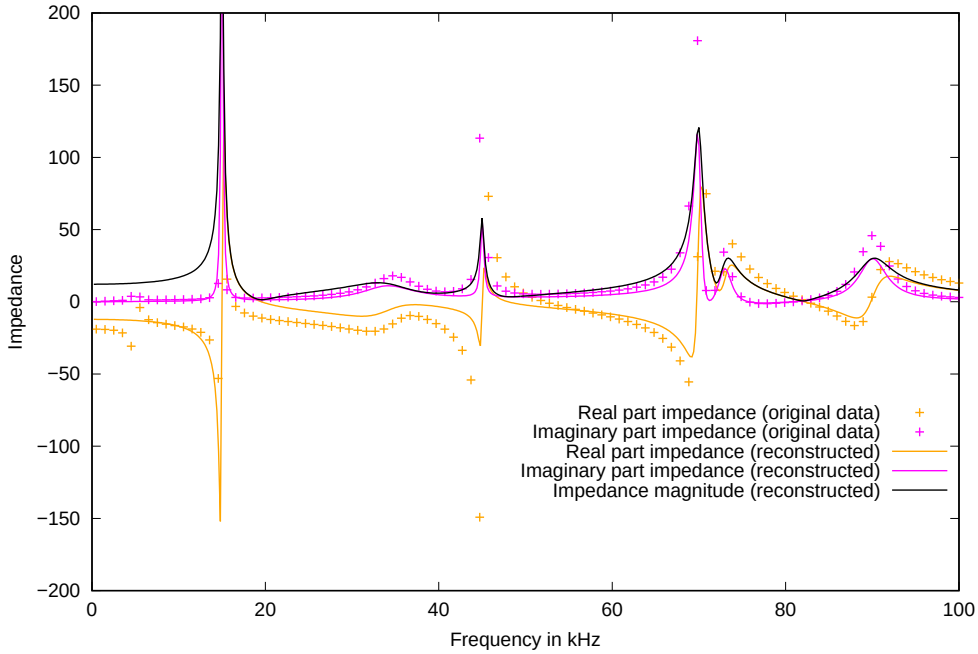

Figure 3-31: *Example 5, Gustavsen's variant 2b for $N = 12$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

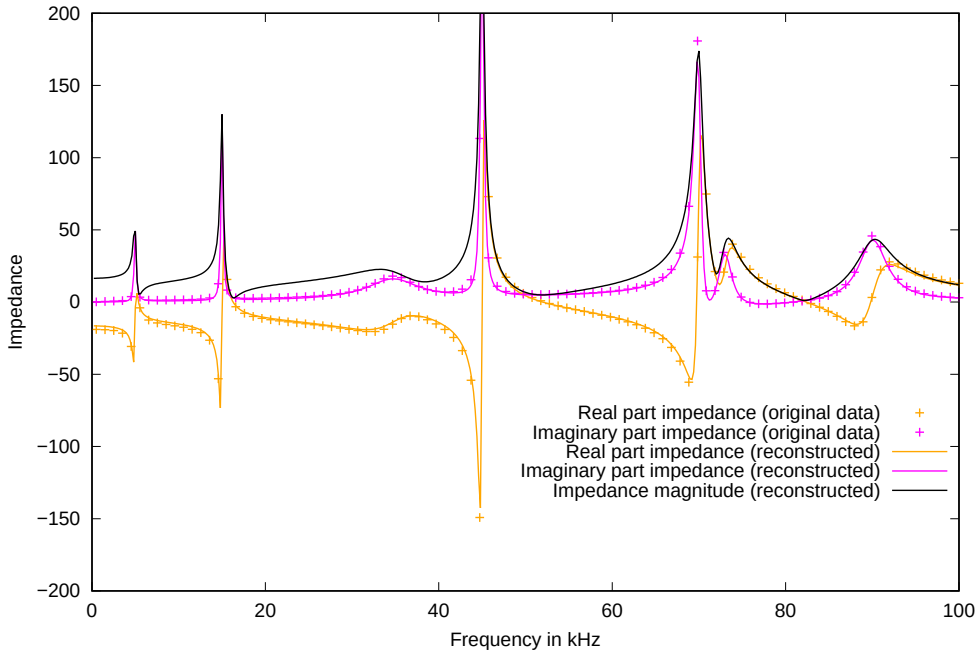

Figure 3-32: *Example 5, Gustavsen's variant 2b for $N = 14$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

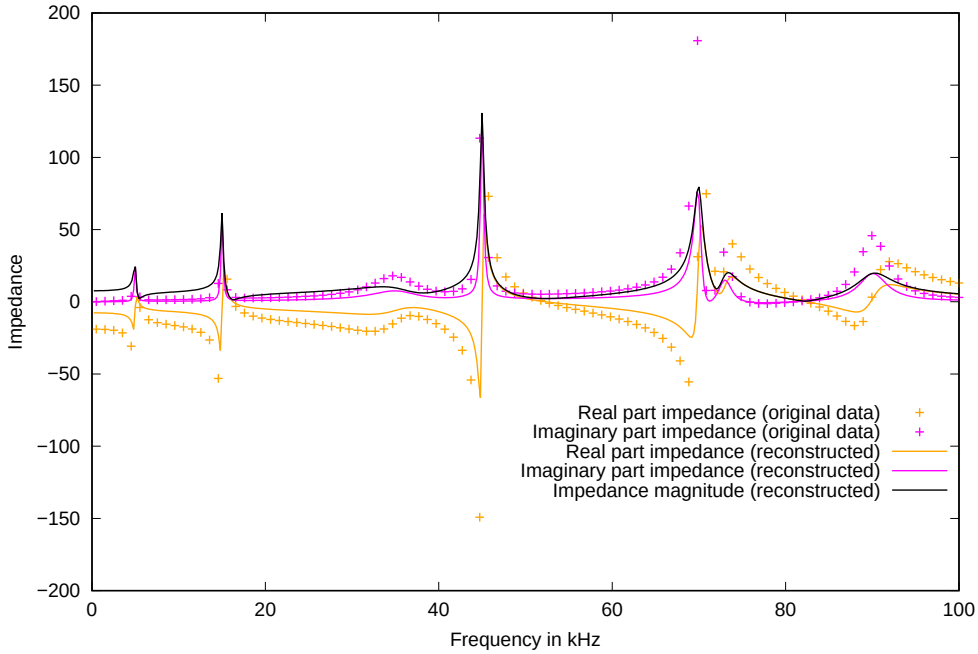

Figure 3-33: *Example 5, Gustavsen's variant 2b for* $N = 16$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

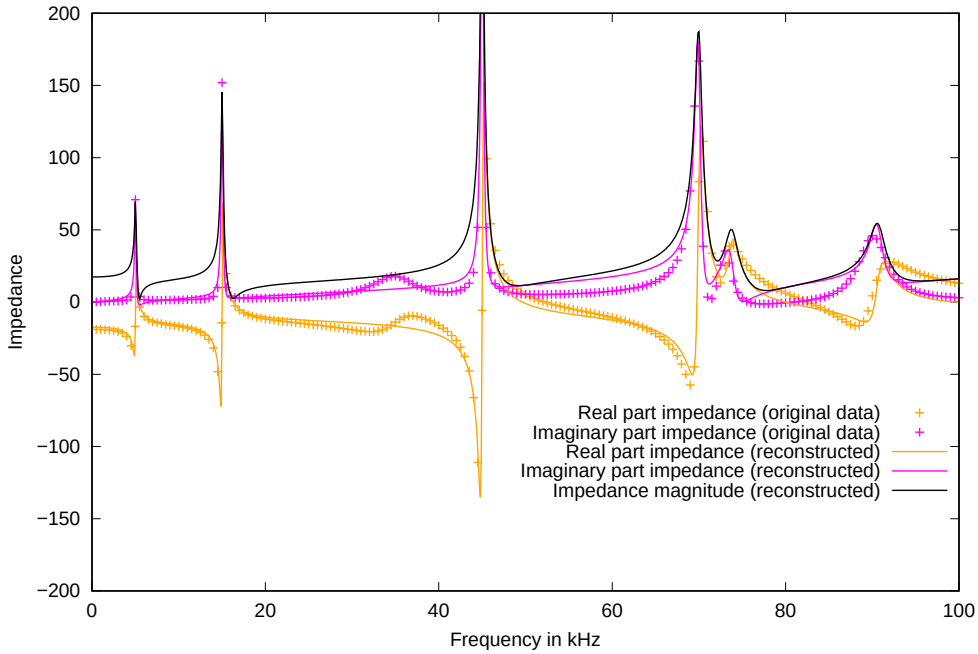

Figure 3-34: *Example 5, Null-space method for* $N = 12$, $L = 10.0$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

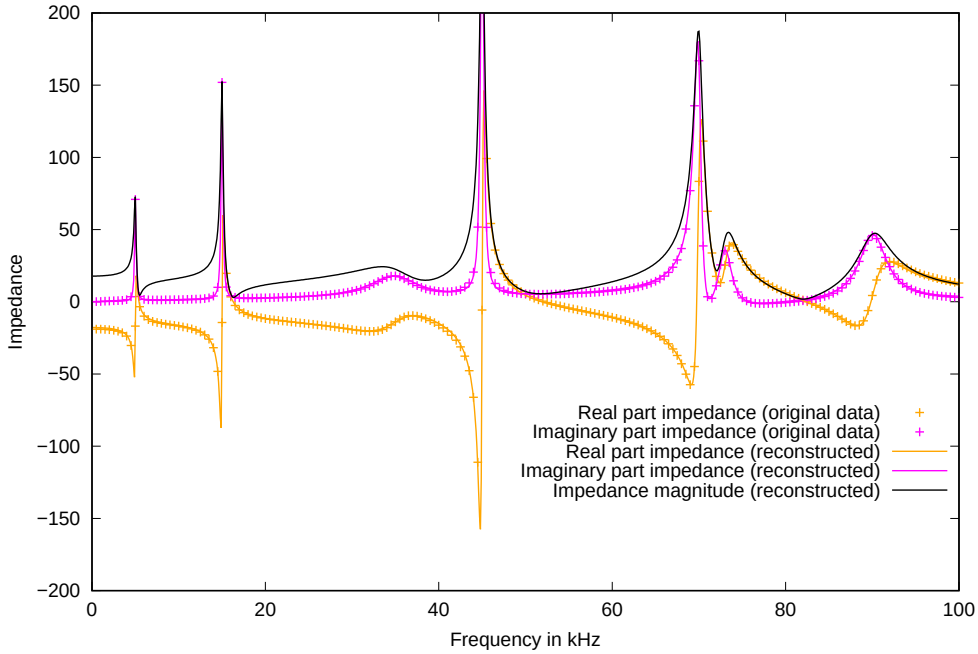

Figure 3-35: *Example 5, Null-space method for $N = 14$, $L = 10.0$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

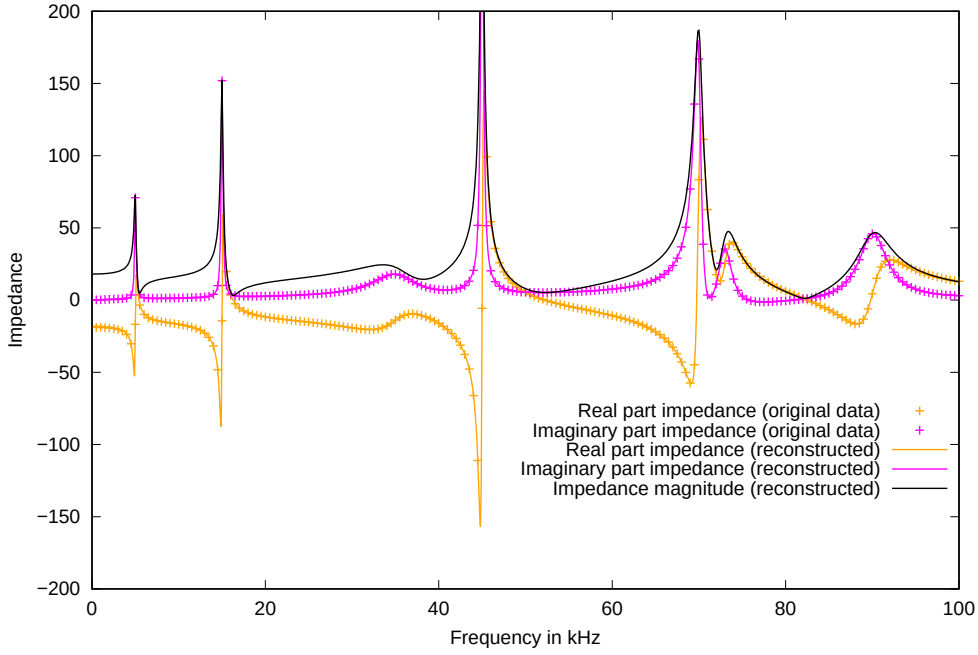

Figure 3-36: *Example 5, Null-space method for $N = 16$, $L = 10.0$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

3.8 Example 6. The 18 pole example with model order overestimation and noise

In this example the previous 18 pole example 3a definitions are employed, but with additive uncorrelated white noise. Noise is added to both the real and imaginary parts of the response so that,

$$Z_{in}(s_i) = \sum_{k=1}^{N} \frac{r_k}{s_i - a_k} + d + e + N_i' + jN_i'' \quad \text{for } i = 1 \text{ to } N_f \tag{3-10}$$

where $s_i = j\omega_i$ and N_i' and N_i'' are real valued noise terms, mutually uncorrelated and uncorrelated with respect to i. Both are taken from a uniform distribution between $-V_n$ to V_n where $V_n = 10.0$. This is similar to example 4.6 of [1], although there the noise was assumed to be real (i.e. $N'' = 0$). The addition of complex noise is likely to provide a better model of a physical noise process. Four variants of the method are considered here; Gustavsen's first variant, the variants 2a and 2b and the null-space method.

Gustavsen's first variant (section 2.7)

When $N = 20$, convergence takes about 15 iterations. The resulting value of $E_2^{(rel)} \approx 0.2$ which coincides with the value of $E_1^{(rel)}$, as shown in figure figure 3-37. After convergence, the RMS errors $E_1^{(rms)} \approx E_2^{(rms)} \approx 7.9$. Note that the RMS value of zero-mean complex noise is defined by,

$$N_{rms} = \sqrt{\langle |z|^2 \rangle}$$

where,

$$z = N_i' + jN_i''$$

and $\langle x \rangle$ represents the expectation value of x. For uniformly distributed uncorrelated complex noise with amplitude V_n this is given by,

$$N_{rms} = \sqrt{\frac{2}{3}} V_n \tag{3-11}$$

so when $V_n = 10.0$, $N_{rms} \approx 8.16$. The converged values for both $E_1^{(rms)}$ and $E_2^{(rms)}$ are thus approximately equal to N_{rms}, which is considered good [2]. However, this is not sufficient for optimal performance since it does not account for the existence of spurious poles with large residues. Figure 3-41 shows the original data and the reconstructed impedance after 50 iterations.

When $N = 40$ convergence is more erratic, as shown in the figure. There are also many more non-negligible spurious resonances predicted. Figure 3-42 shows the original data and the reconstructed impedance after 50 iterations.

Gustavsen's variant 2a (section 2.8)

When $N = 20$, convergence to the expected value (that predicted above) does not occur, with $E_2^{(rel)} \sim 1$. On the other hand $E_1^{(rel)}$ oscillates with values between about 0.1 and 0.01 as illustrated in figure 3-38. This is accompanied by a similar oscillation in the predicted value

for $\tilde{d}$ (not shown). Interestingly, the condition number for the matrix Ω remains acceptable with its rank remaining equal to its dimension. Figure 3-43 shows the original data and the reconstructed impedance after 50 iterations, showing that the algorithm badly fails.

When $N = 40$, performance is similarly poor with errors as shown in the above figure. As above, the matrix Ω remains acceptable with its rank remaining equal to its dimension (in this case taking the value 82) so there is no loss of information when determining the residues and value for $\tilde{h}$. The poor performance can therefore not be attributed to the use of the Moore-Penrose inverse. Figure 3-44 shows the original data and the reconstructed impedance after 50 iterations.

Gustavsen's variant 2b (section 2.8)

When $N = 20$, convergence to the expected value fails to occur, with $E_2^{(rel)} \sim 1$. The measure $E_1^{(rel)}$ oscillates wildly with quite small values for a while before stabilising. This is shown in figure 3-39. There is a similar oscillation in the predicted value for $\tilde{d}$ (not shown) stabilising at the value $\tilde{d} \sim 0.187$. For a proper convergence $\tilde{d} \sim 1$. Figure 3-45 shows the original data and the reconstructed impedance after 50 iterations.

When $N = 40$, there is a similar lack of convergence as shown in the figure. The poor performance is shown in the reconstructed impedance, figure 3-46.

Null-space method (section 2.6)

When $N = 20$, 9 eigenvectors are automatically selected to form the null-space. After about 10 iterations the relative error $E_2^{(rel)} \sim 0.2$. After 50 iterations, the RMS errors $E_1^{(rms)} \approx E_2^{(rms)} \approx 8.1$ similar to use of Gustavsen's first variant. See figure 3-40. The noise prevents accurate determination of the pole positions but the reconstructed impedance is good, as shown in figure 3-47. Here, no spurious poles are predicted with large residues.

When $N = 40$, 17 eigenvalues are automatically selected to form the null-space. After about 5 iterations the relative error $E_2^{(rel)} \sim 0.2$. The errors then behave as above, as shown in the figure. The reconstructed impedance shows a number of spurious resonances, but the original function is adequately reproduced, as shown in figure 3-48.

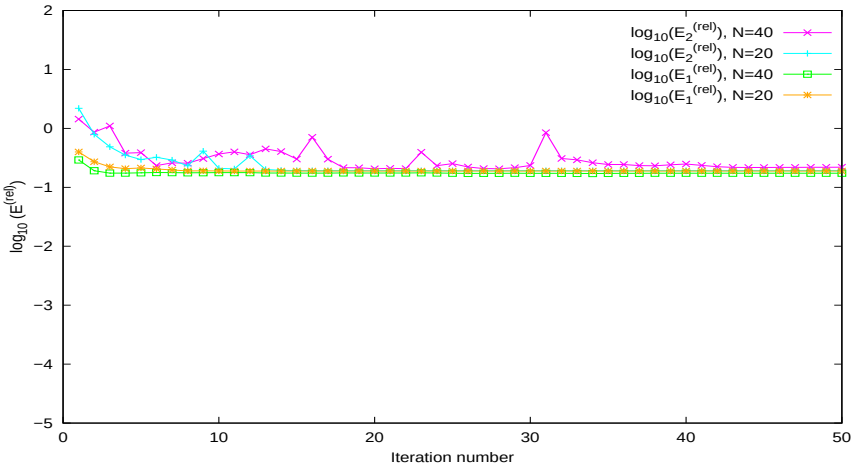

Figure 3-37: *Example 6, Gustavsen's first variant.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

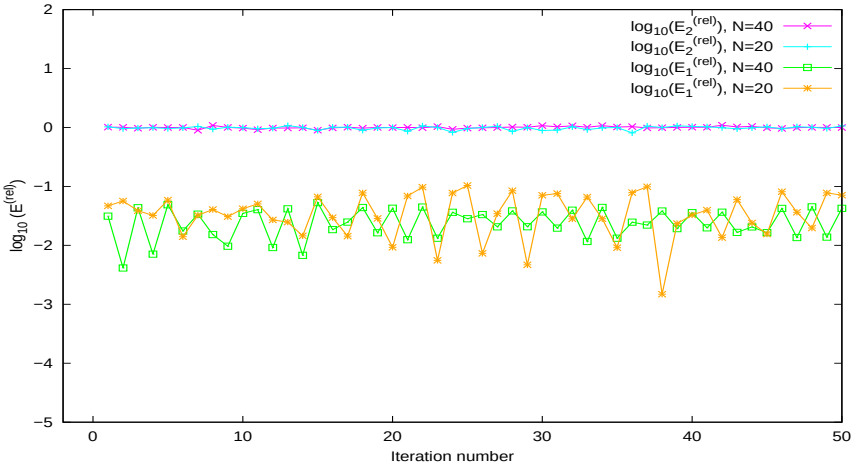

Figure 3-38: *Example 6, Gustavsen's variant 2a.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

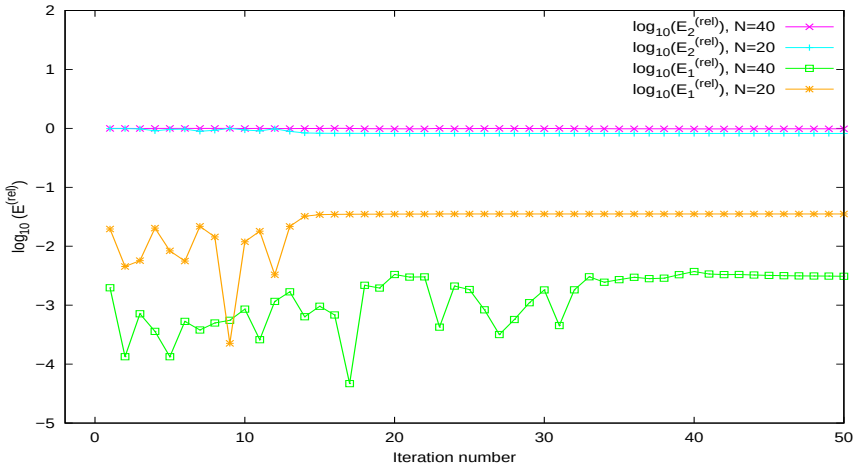

Figure 3-39: *Example 6, Gustavsen's variant 2b.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

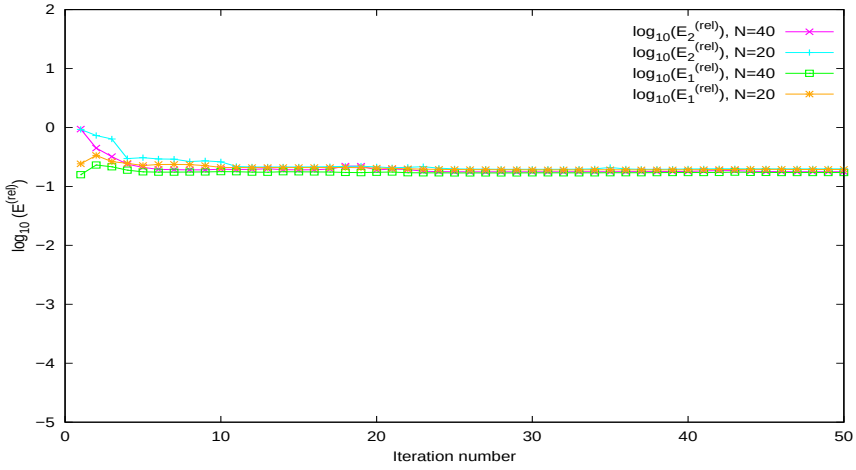

Figure 3-40: *Example 6, Null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *for* $N = 20$ *and* $N = 40$ *vs iteration count*

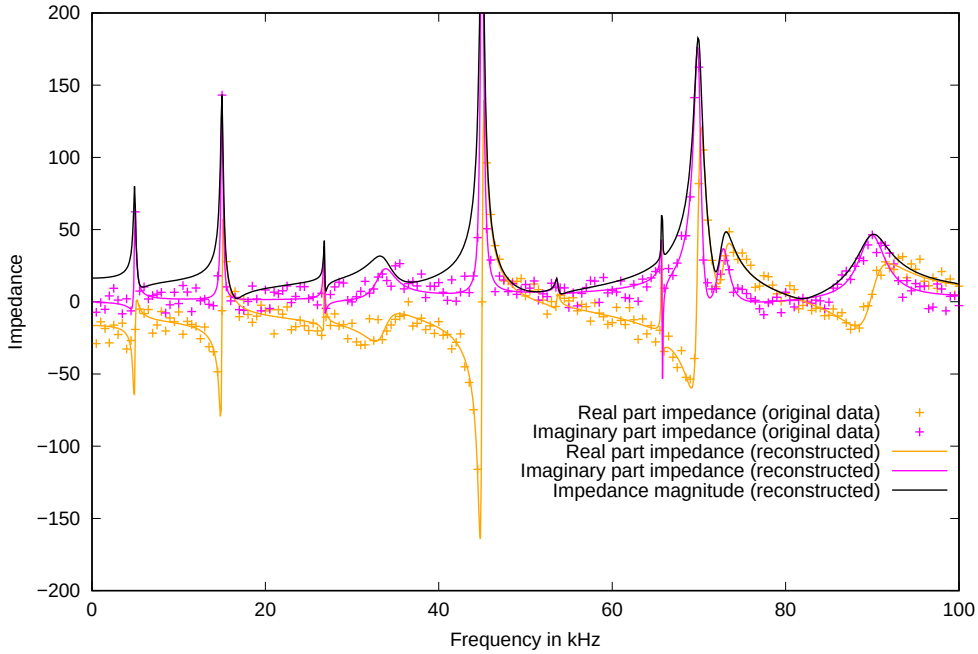

Figure 3-41: *Example 6, Gustavsen's first variant for* $N = 20$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

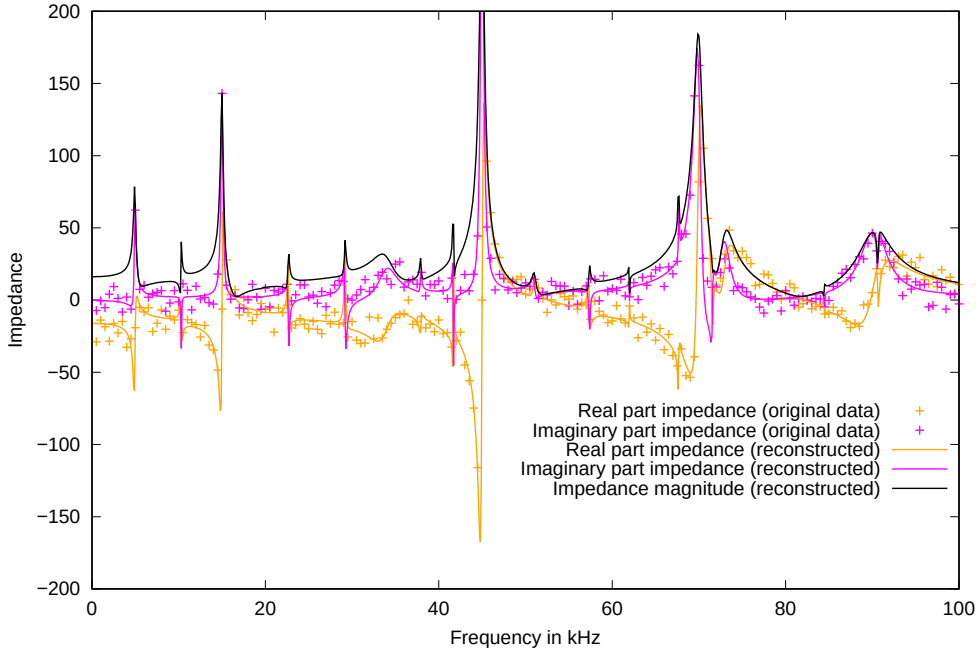

Figure 3-42: *Example 6, Gustavsen's first variant for* $N = 40$. *Impedance function fit with* Z_{in} *as a function of frequency in kHz.*

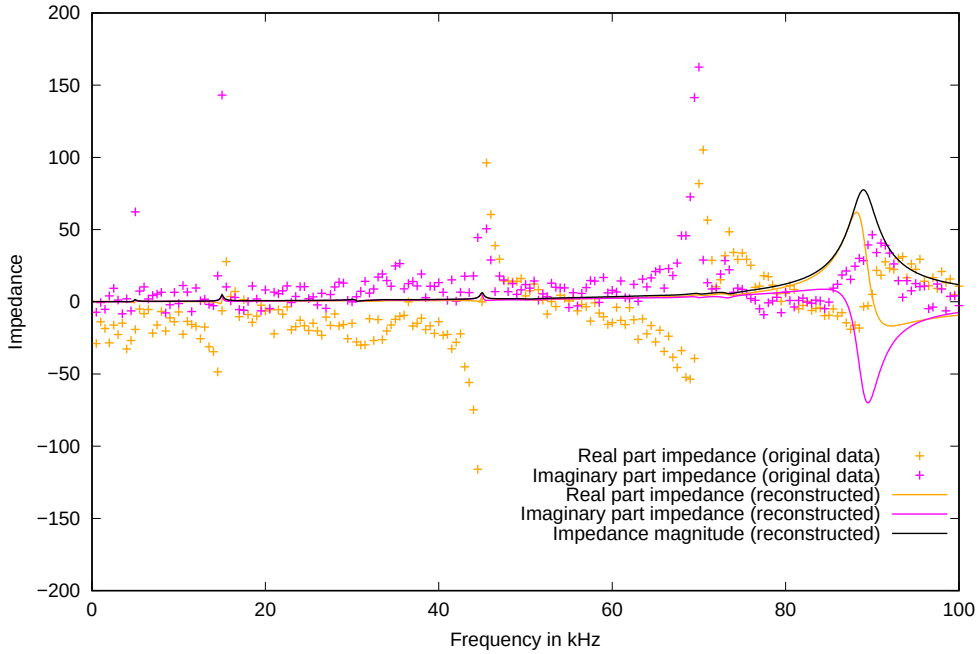

Figure 3-43: *Example 6, Gustavsen's variant 2a for $N = 20$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

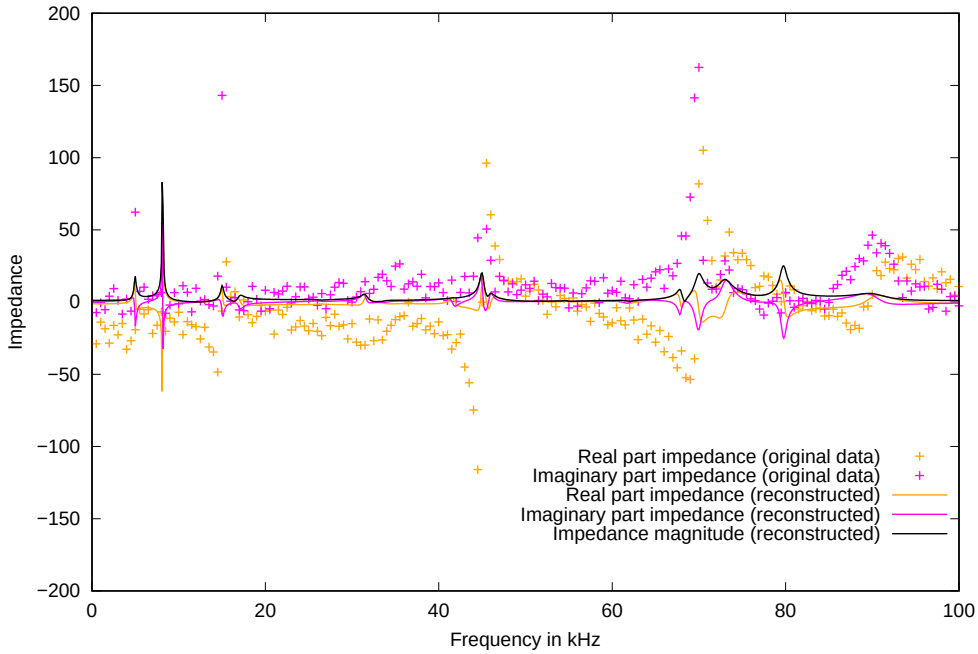

Figure 3-44: *Example 6, Gustavsen's variant 2a for $N = 40$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

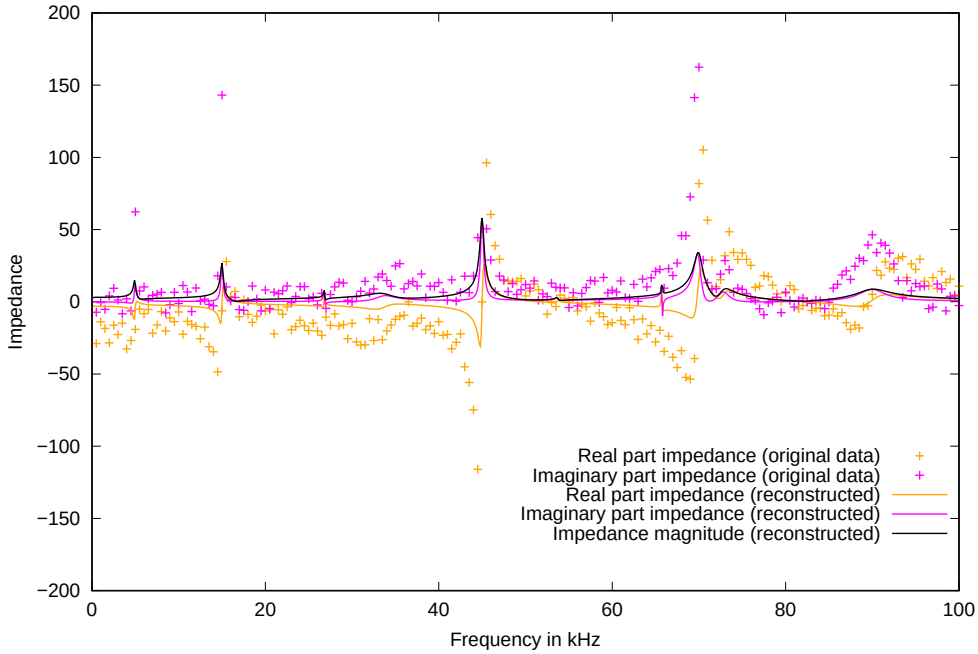

Figure 3-45: *Example 6, Gustavsen's variant 2b for $N = 20$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

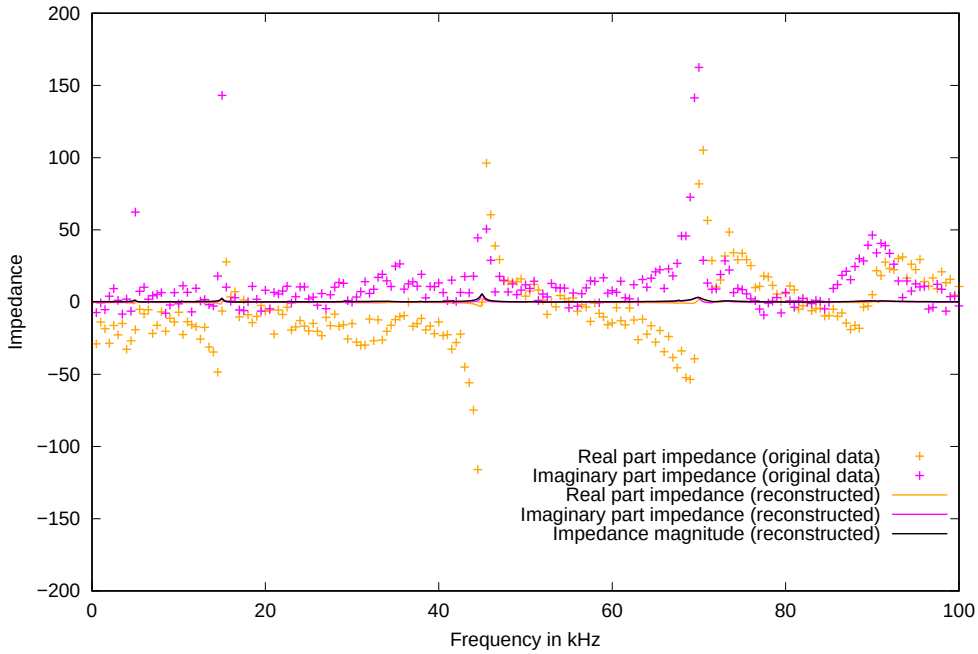

Figure 3-46: *Example 6, Gustavsen's variant 2b for $N = 40$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

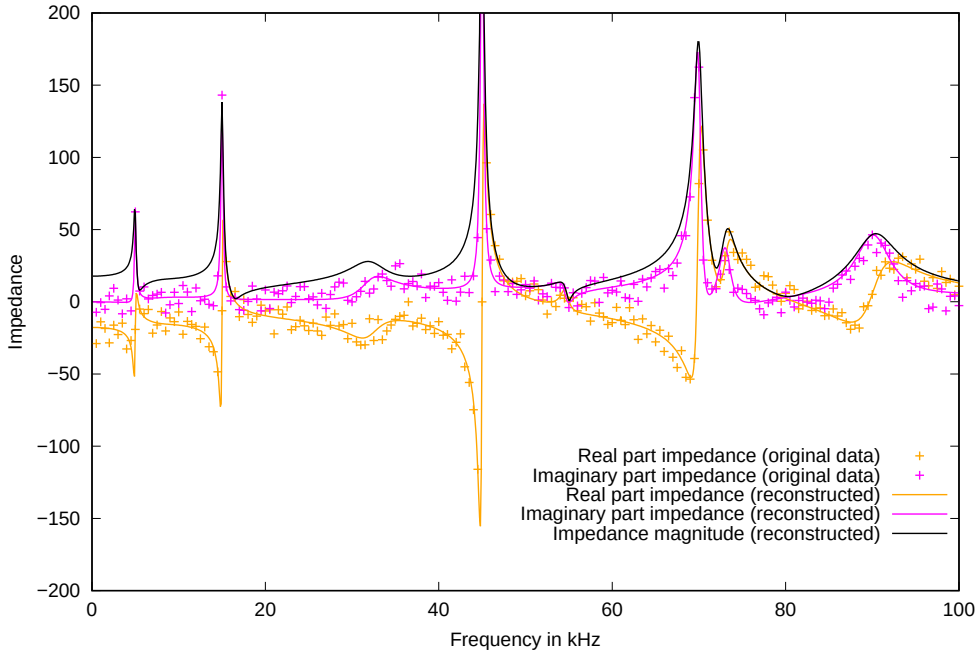

Figure 3-47: *Example 6, Null-space method for $N = 20$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

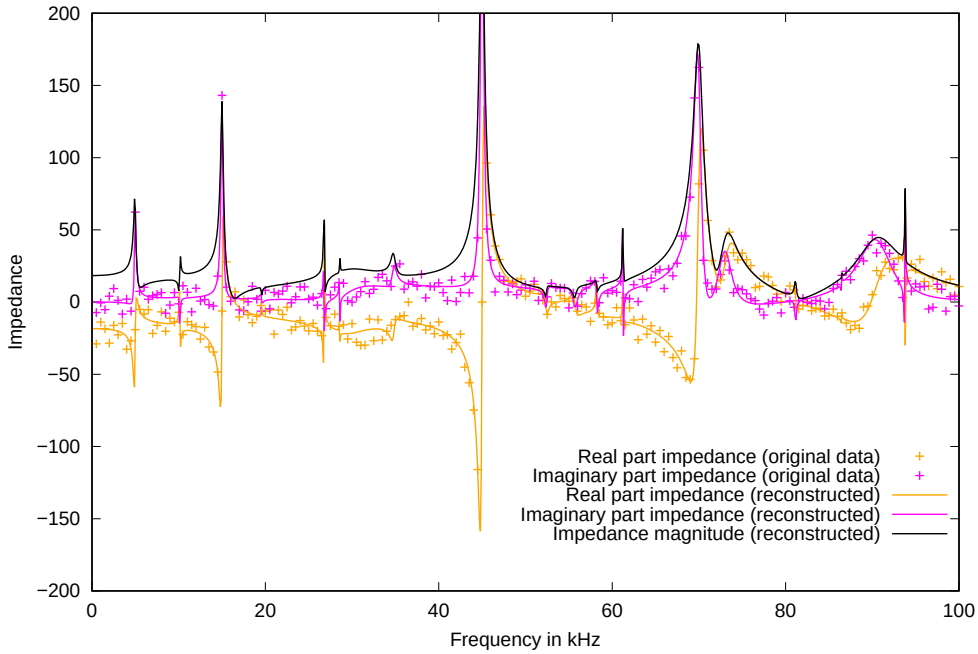

Figure 3-48: *Example 6, Null-space method for $N = 40$. Impedance function fit with Z_{in} as a function of frequency in kHz.*

3.9 Interim conclusions and application to radar absorbent materials

The previous tests all assume an impedance function of finite order. Under this assumption the null-space method consistently performs well and is generally robust provided the model order is not significantly underestimated. When the model order is correctly specified and no noise is present, accuracy is generally highest when using Gustavsen's first variant and in some cases the second variants. When the model order is overestimated with no noise present, both Gustavsen's first and second variants lack robustness and the null-space method is favoured. When the model order is overestimated and there is significant noise, Gustavsen's first variant and the null-space method perform well but Gustavsen's second variants perform poorly. When the model order is underestimated, Gustavsen's first method perform well and the null-space method can do so but is less robust with accuracy that depends on the parameter L.

In the next section a real-world problem is considered. Two of the algorithms, Gustavsen's first variant and the null-space method, are applied to characterise the performance of a radar absorbent material (RAM) at normal incidence. This assumes that the complex reflection coefficient is accurately determined. The eventual aim is to determine how the positions of the poles change as a function of angle of incidence, but this will not be reported here. The RAM problem is a class of transmission line problem where loss is large over a wide bandwidth. Consequently the impedance function is of infinite order with poles that may have both small and large real parts in the complex plane. This presents a challenge to all pole determination algorithms.

Given that the complex reflection coefficient is known, an impedance function is constructed according to (2-1). The particular structure to be considered is defined in [7] and refers to an ultra-wide band RAM of generalised Jaumann type. The particular absorber described is shown in figure 3-49 and consists of five layers of lossless dielectric backed by a perfect conductor. At the interface between layers 1 and 2 (layer 1 counted nearest the perfect conductor) is a thin resistive material embedded in a thin dielectric. At the interfaces between layers 2 and 3, and 3 and 4 is a metasurface embedded in a thin dielectric to lend mechanical integrity. The metasurfaces are represented as zero-thickness surface impedance sheets whose equivalent circuit is represented by a series resistance and capacitance (RC). Practical techniques for the synthesis of such metasurfaces are described on [7]. These kinds of radar absorbers are amongst the widest bandwidth RAM known to exist which are also capable of meeting the Rozanov/Brewitt-Taylor bandwidth/thickness maximum performance bounds [8], [9], [10].

Figure 3-50 shows the predicted reflection coefficient magnitude, in dB, for a plane wave at normal incidence and at 45 degrees for the two principal polarisations (TE and TM), over the frequency range 0 to 40 GHz. [2] In the first part of the study Gustavsen's first variant,

[2] Note that the predictions differ slightly from [7] because in the original report RC values for this particular geometry were not given, only the metasurface structure details. Here, $R_1 = 180\Omega$, $R_2 = 450\,\Omega$, $C_2 = 0.084$ pF, $R_3 = 720\,\Omega$, $C_3 = 0.060$ pF.

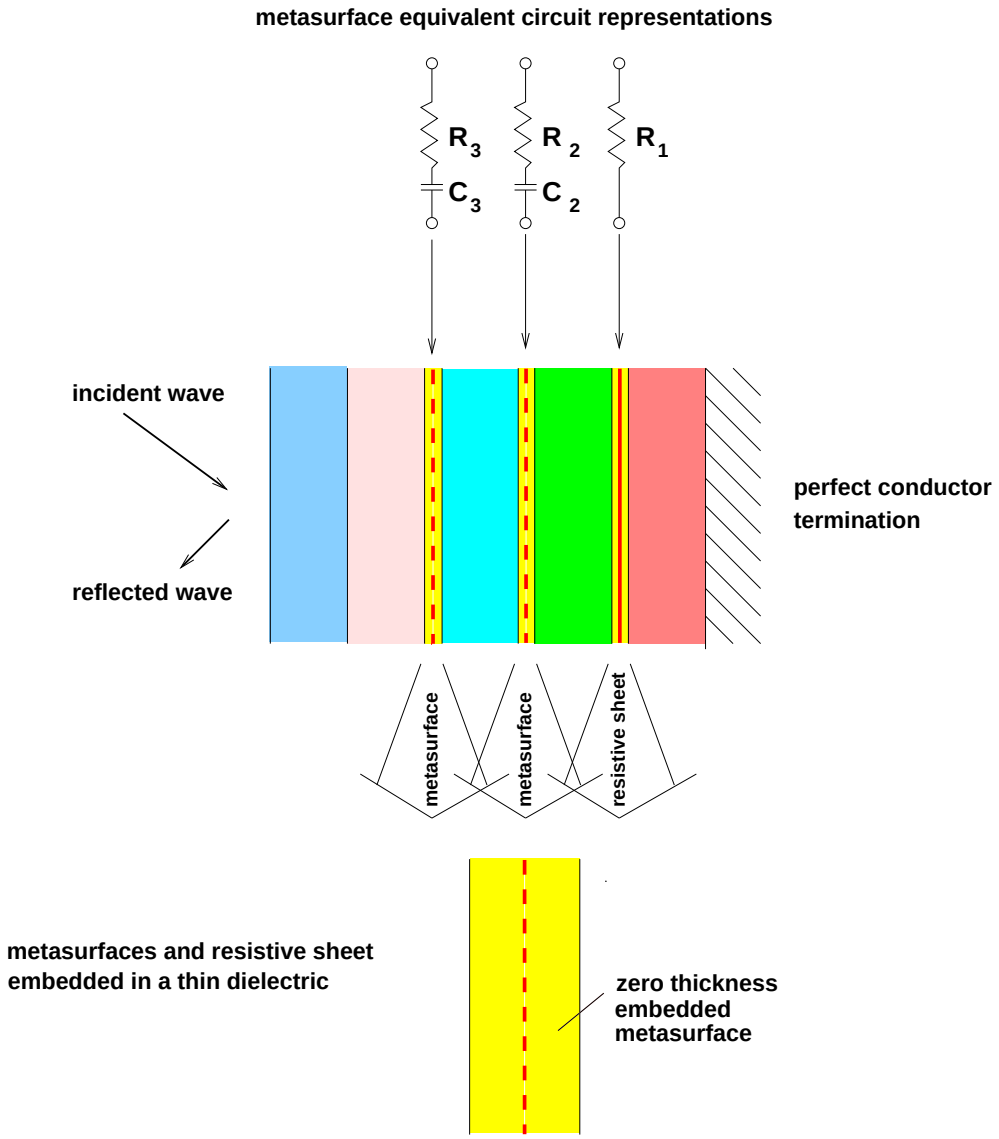

Figure 3-49: *Kazemzadeh and Karlsson's generalised Jaumann radar absorbing structure with five layers and embedded metasurfaces*

the variant 2a and the null-space method are employed in a characterisation of the input impedance, normalised to the characteristic impedance of free space, at normal incidence. A model will be constructed assuming the model order $N = 8, 10, 12, 14, 16, 18$ and 24 with d and e parameters required. All frequencies are normalised to the upper frequency $\omega_{sc} = 2\pi f_{max}$ where $f_{max} = 40$ GHz. Sampling is uniform and between 0.05 GHz and 40.00 GHz in 0.05 GHz intervals, so $N_f = 800$.

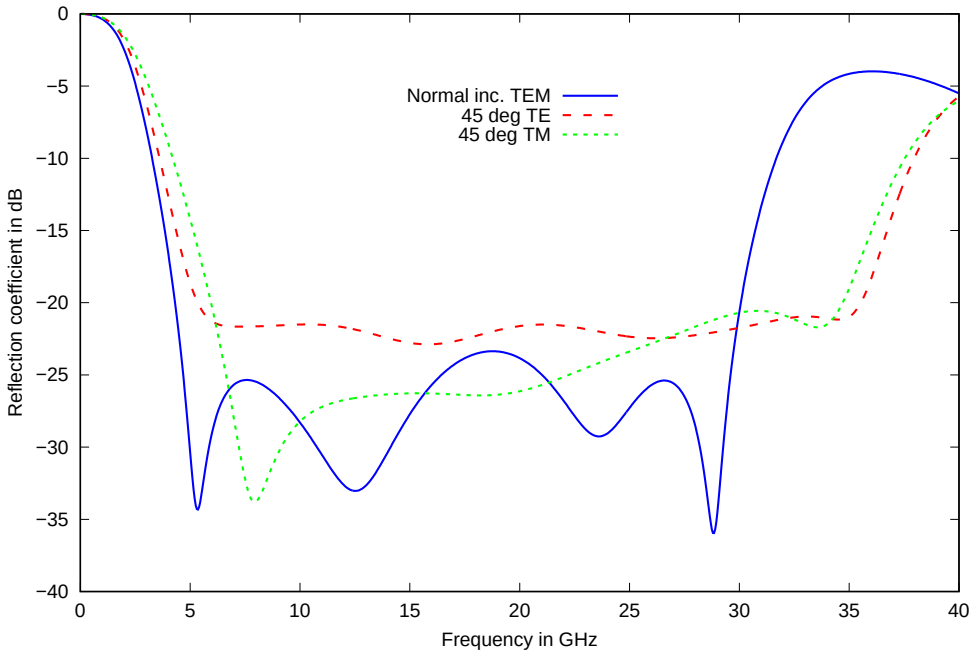

Figure 3-50: *Predicted reflection coefficient as a function of frequency for the Kazemzadeh-Karlsson structure*

For this study the null-space method assumes a value of $L = 1000$, with no attempt made to reduce the fitting error $E_2^{(rel)}$ by changing its value. Figures 3-51 to 3-57 plot the log of the relative errors, $\log_{10}(E_1^{(rel)})$ and $\log_{10}(E_2^{(rel)})$, as a function of iteration number for assumed model orders $N = 8$ to $N = 24$, respectively. Each figure shows the predicted error using Gustavsen's first variant and the null-space method. It is apparent that for $N > 16$ Gustavsen's first variant completely fails to reduce the error $E_2^{(rel)}$ to the expected value between 10^{-5} and 10^{-6} whereas the null-space method is accurate here. When $N = 16$, Gustavsen's first variant fails and the null-space method is not very accurate. When $N = 14$, Gustavsen's first variant and the null-space method are both accurate and convergent, although Gustavsen's method is more accurate. For $N < 14$ both methods are convergent but Gustavsen's method becomes much more accurate and the null-space method is unsuitable. These results are in line with the findings of the previous more contrived examples. The poles and residues predicted with the null-space method assuming $N = 18$ (one shown for each complex conjugate pair) are tabulated below:

pair designator $N = 18$	pole positions in GHz (real, imaginary)	residues in GHz (real, imaginary)
1	(-0.1064223E+02 , 0.5713983E+02)	(0.8868187E+01 , 0.9349189E+01)
2	(-0.1317163E+01 , 0.4151803E+02)	(0.3442899E+01 , 0.3093899E+00)
3	(-0.1941805E+01 , 0.3116026E+02)	(0.1308345E+01 , -0.9399161E+00)
4	(-0.6102873E+01 , 0.2475480E+02)	(0.2856192E+01 , -0.8493799E+00)
5	(-0.3717194E+01 , 0.1982150E+02)	(-0.2482107E-02 , -0.4194086E-02)
6	(-0.7279361E+01 , 0.1455122E+02)	(0.2162545E+01 , -0.8013764E+00)
7	(-0.9133138E+00 , 0.2879012E+01)	(0.9163199E+00 , 0.9865479E+00)
8	(-0.4242769E+01 , 0.1002802E+02)	(-0.2173924E-01 , -0.1347671E-01)
9	(-0.6599745E+01 , 0.7154632E+01)	(0.1559405E+01 , 0.1696901E+01)
	d= 0.3565236E+00	e=0.2610644E-02

The poles and zeros predicted with Gustavsen's first variant assuming $N = 14$ are tabulated as:

pair designator $N = 14$	pole positions in GHz (real, imaginary)	residues in GHz (real, imaginary)
1	(-0.9945951E+01 , 0.5894121E+02)	(0.1110095E+02 , 0.8074493E+01)
2	(-0.1319784E+01 , 0.4151862E+02)	(0.3450940E+01 , 0.3186048E+00)
3	(-0.1941535E+01 , 0.3115995E+02)	(0.1308313E+01 , -0.9389356E+00)
4	(-0.6096706E+01 , 0.2470874E+02)	(0.2888710E+01 , -0.7603347E+00)
5	(-0.7396987E+01 , 0.1425796E+02)	(0.2544779E+01 , -0.5126336E+00)
6	(-0.6436867E+01 , 0.7288122E+01)	(0.1673775E+01 , 0.1385921E+01)
7	(-0.9133100E+00 , 0.2878989E+01)	(0.9162463E+00 , 0.9865864E+00)
	d= 0.2594494E+00	e=0.3199958E-02

These results confirm the requirement for 7 dominant pole pairs. The pole pairs 5 and 8 may be ignored in the $N = 18$ simulation with the remaining poles in good agreement. Interestingly, both sets of results require a pole pair between 57 GHz and 59 GHz, well outside the data span.

The regenerated normalised impedances and reflection coefficient magnitudes (in dB) are shown for $N = 12, 14$ and 18 under the two methods. Figures 3-58 and 3-59 show the regenerated impedance (normalised to free space) and the reflection coefficient, respectively for the $N = 12$ Gustavsen variant 1. Figures 3-60 and 3-61 show the regenerated impedance (normalised to free space) and the reflection coefficient, respectively for the $N = 12$ null-space method. Figures 3-62 and 3-63 show the regenerated impedance (normalised to free space) and the reflection coefficient, respectively for the $N = 14$ Gustavsen variant 1. Figures 3-64 and 3-65 show the regenerated impedance (normalised to free space) and the reflection coefficient, respectively for the $N = 14$ null-space method. Figures 3-66 and 3-67 show the

regenerated impedance (normalised to free space) and the reflection coefficient, respectively for the $N = 18$ Gustavsen variant 1. Figures 3-68 and 3-69 show the regenerated impedance (normalised to free space) and the reflection coefficient, respectively for the $N = 18$ null-space method.

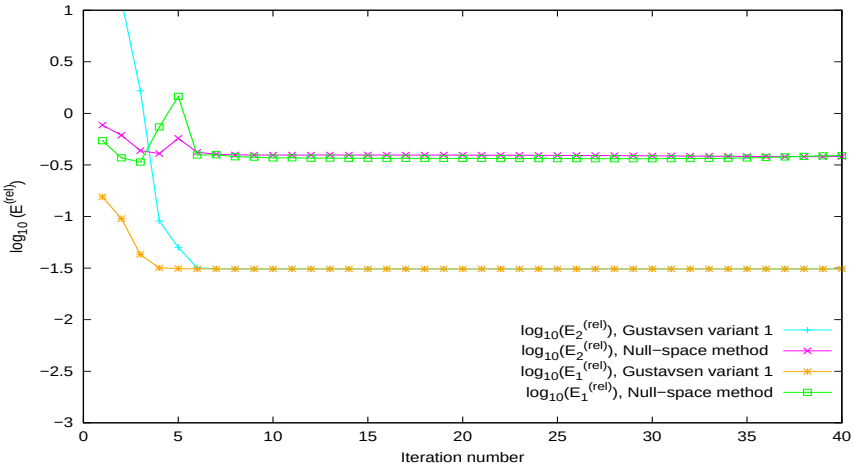

Figure 3-51: *Example 7, Gustavsen's first variant and null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration number for* $N = 8$

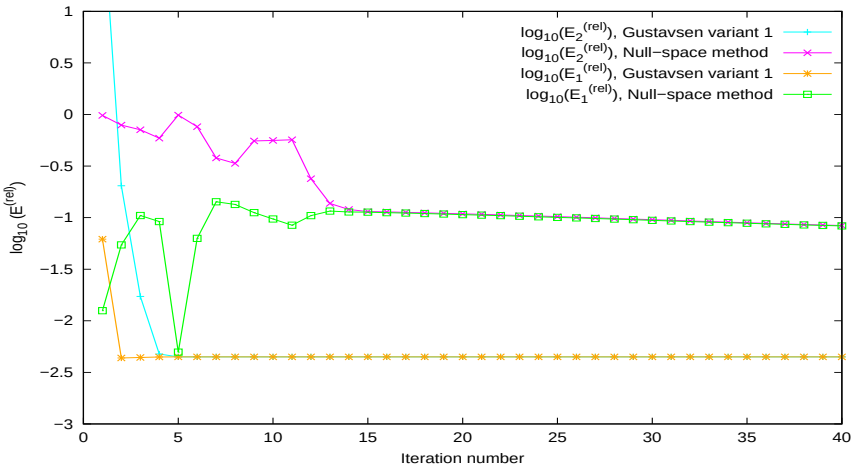

Figure 3-52: *Example 7, Gustavsen's first variant and null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration number for* $N = 10$

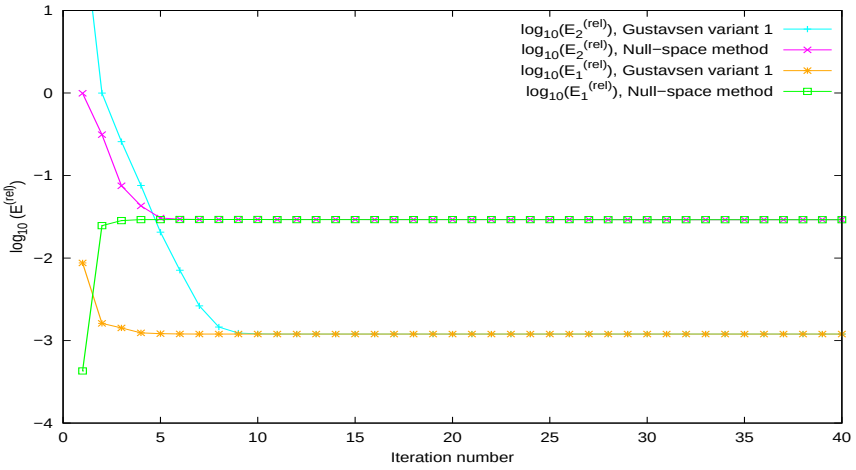

Figure 3-53: *Example 7, Gustavsen's first variant and null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration number for* $N = 12$

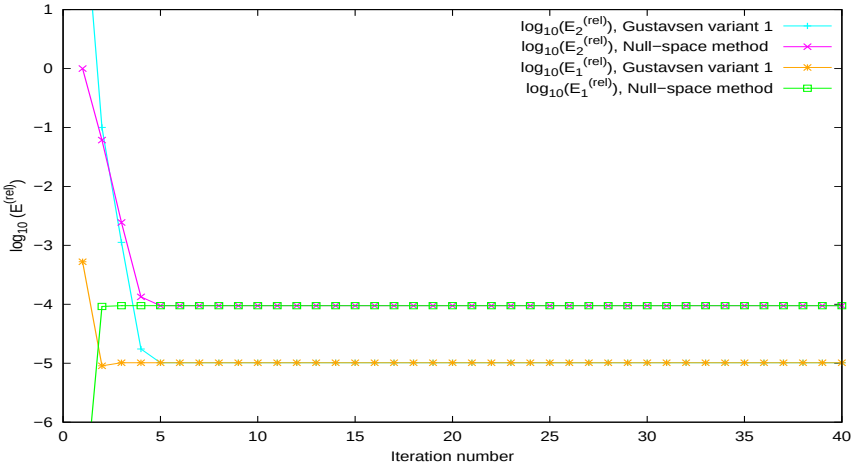

Figure 3-54: *Example 7, Gustavsen's first variant and null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration number for* $N = 14$

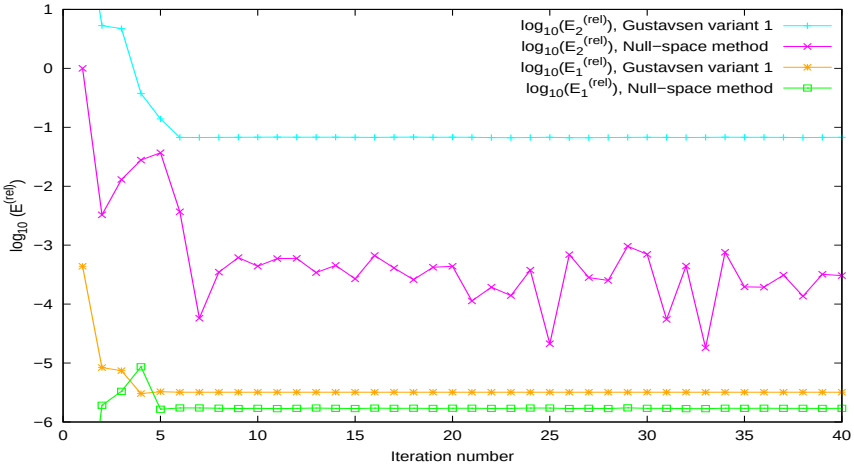

Figure 3-55: *Example 7, Gustavsen's first variant and null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration number for* $N = 16$

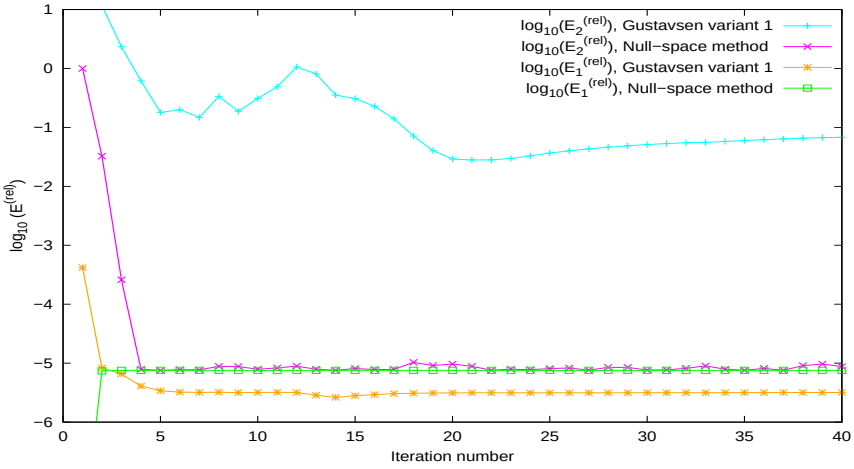

Figure 3-56: *Example 7, Gustavsen's first variant and null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration number for* $N = 18$

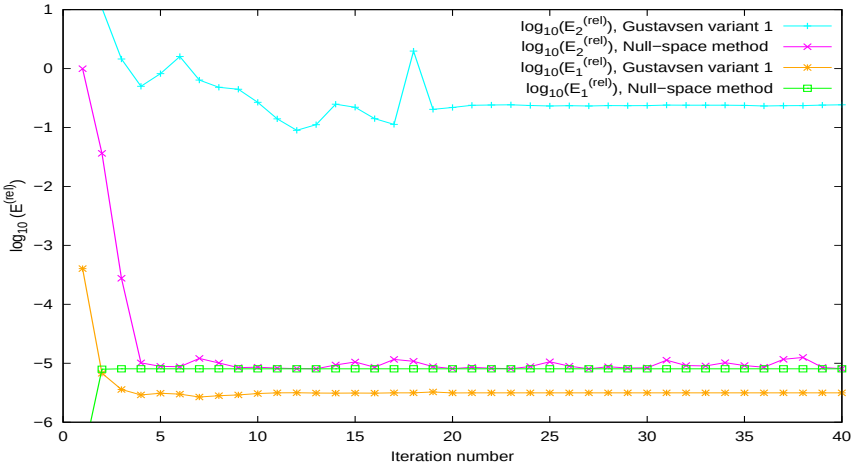

Figure 3-57: *Example 7, Gustavsen's first variant and null-space method.* $E_1^{(rel)}$ *and* $E_2^{(rel)}$ *vs iteration number for* $N = 24$

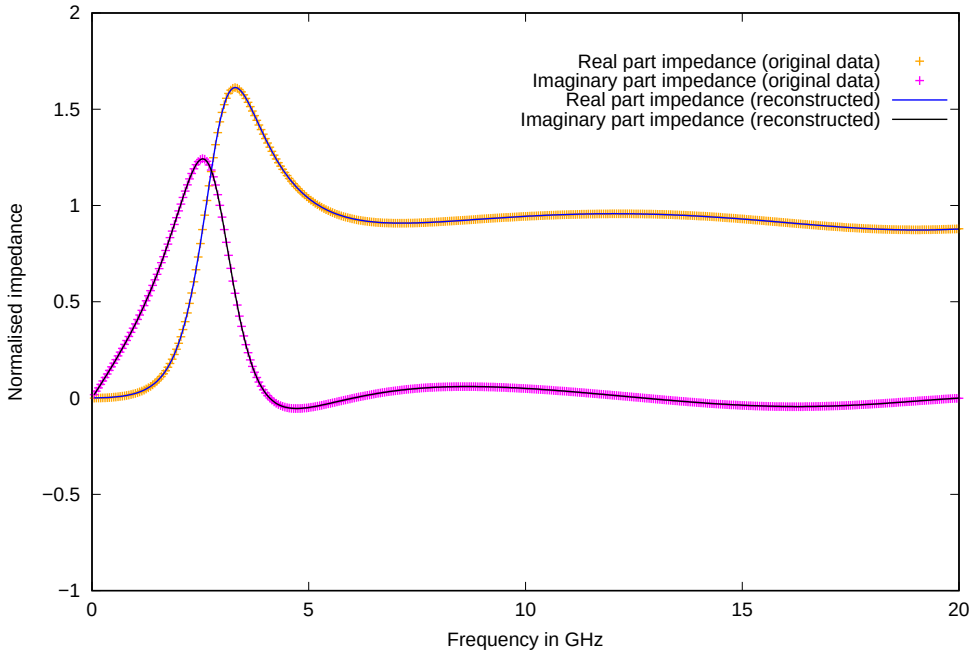

Figure 3-58: *Example 7, Gustavsen variant 1, for $N = 12$. Impedance function fit with Z_{in} as a function of frequency in GHz.*

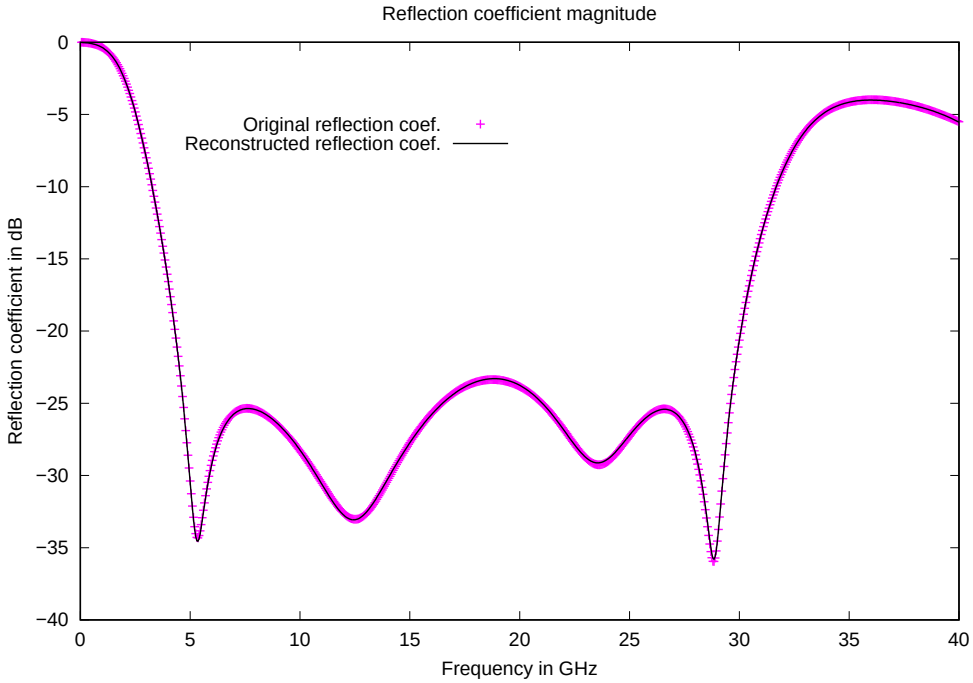

Figure 3-59: *Example 7, Gustavsen variant 1, for $N = 12$. Reflection coefficient magnitude in dB as a function of frequency in GHz.*

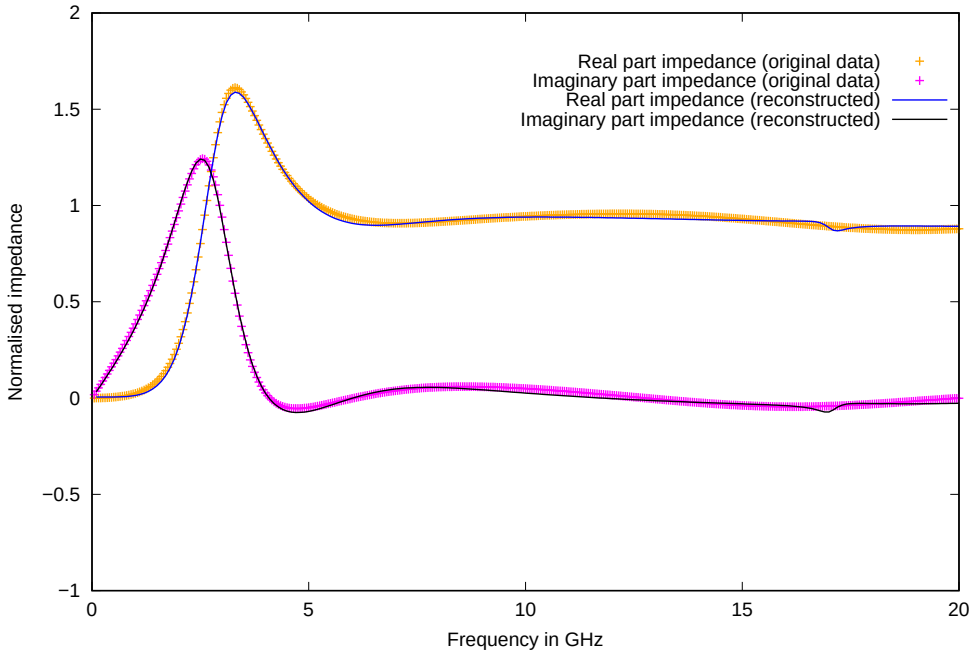

Figure 3-60: *Example 7, Null-space method, for $N = 12$. Impedance function fit with Z_{in} as a function of frequency in GHz.*

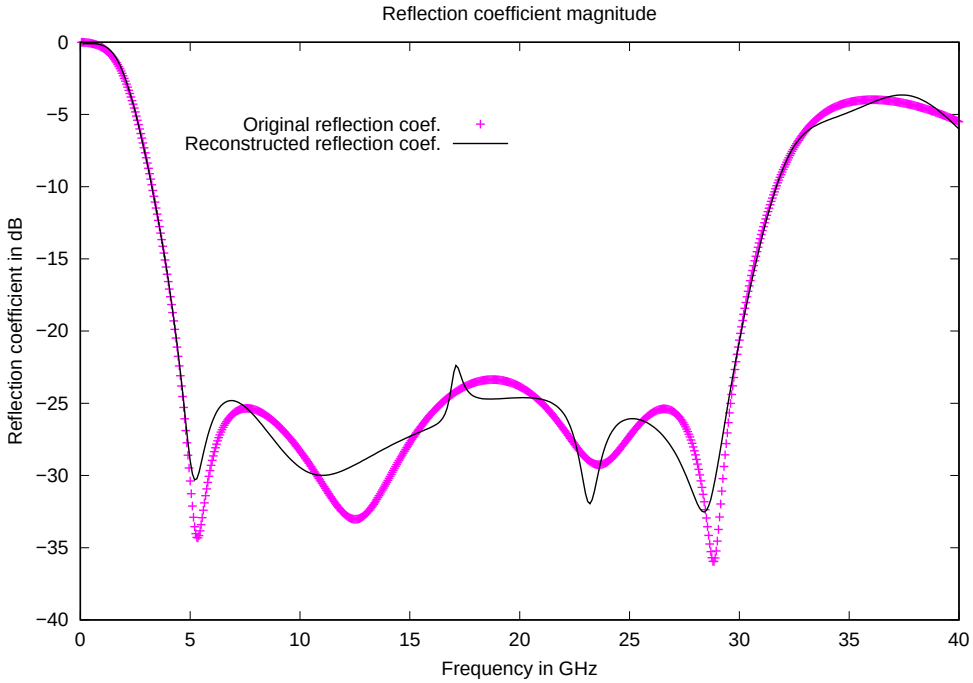

Figure 3-61: *Example 7, Null-space method, for $N = 12$. Reflection coefficient magnitude in dB as a function of frequency in GHz.*

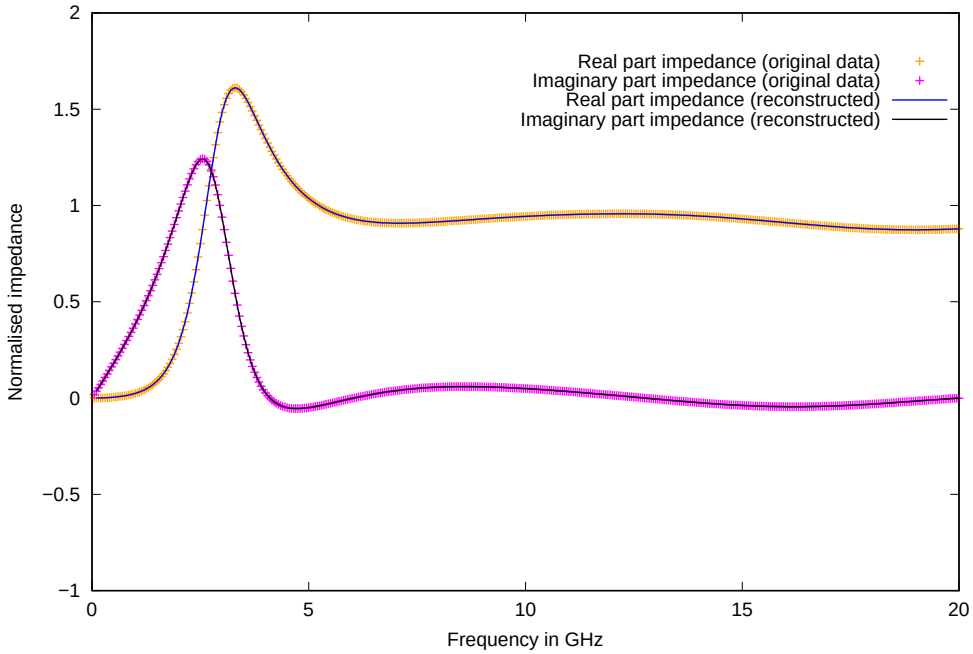

Figure 3-62: *Example 7, Gustavsen variant 1, for $N = 14$. Impedance function fit with Z_{in} as a function of frequency in GHz.*

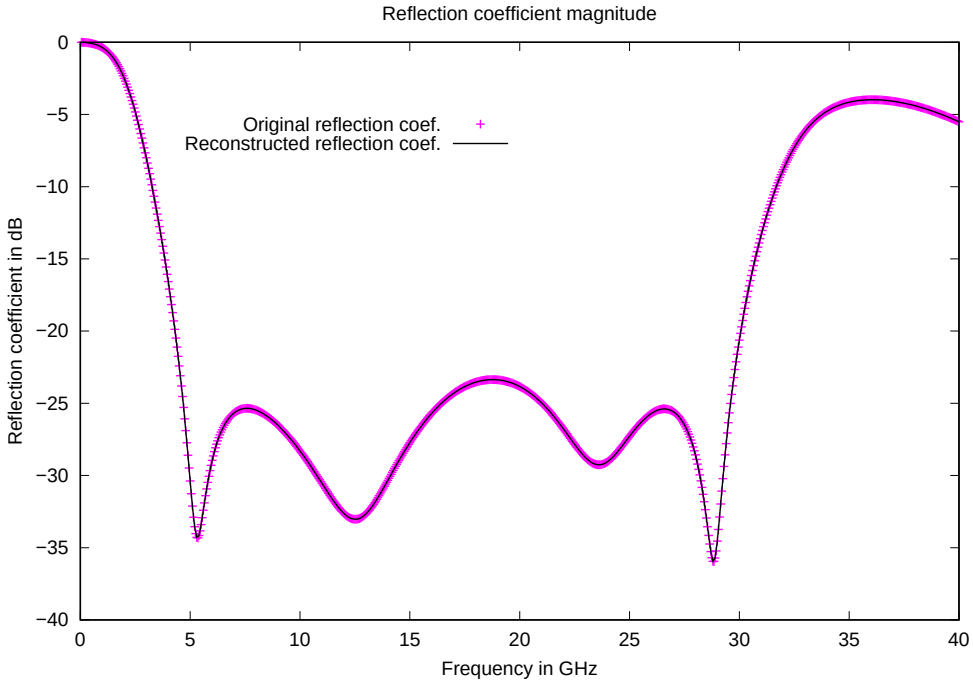

Figure 3-63: *Example 7, Gustavsen variant 1, for $N = 14$. Reflection coefficient magnitude in dB as a function of frequency in GHz.*

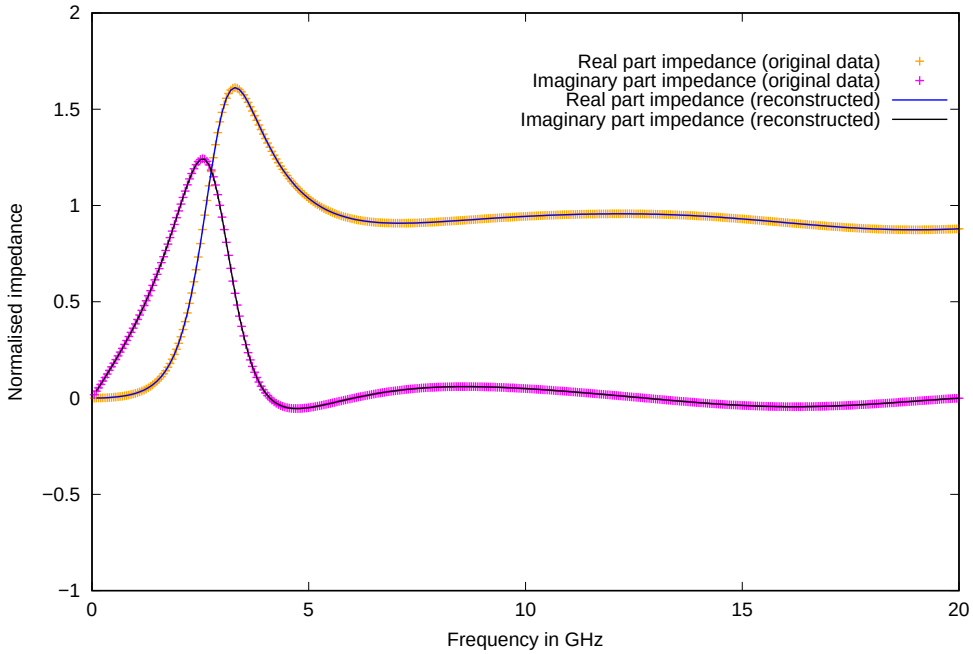

Figure 3-64: *Example 7, Null-space method, for $N = 14$. Impedance function fit with Z_{in} as a function of frequency in GHz.*

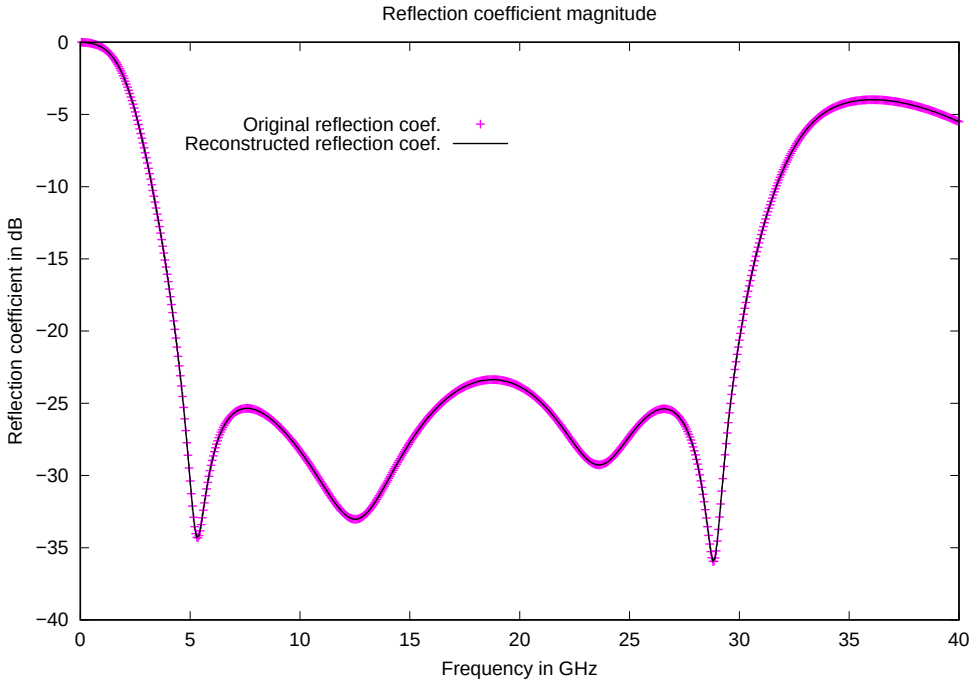

Figure 3-65: *Example 7, Null-space method, for $N = 14$. Reflection coefficient magnitude in dB as a function of frequency in GHz.*

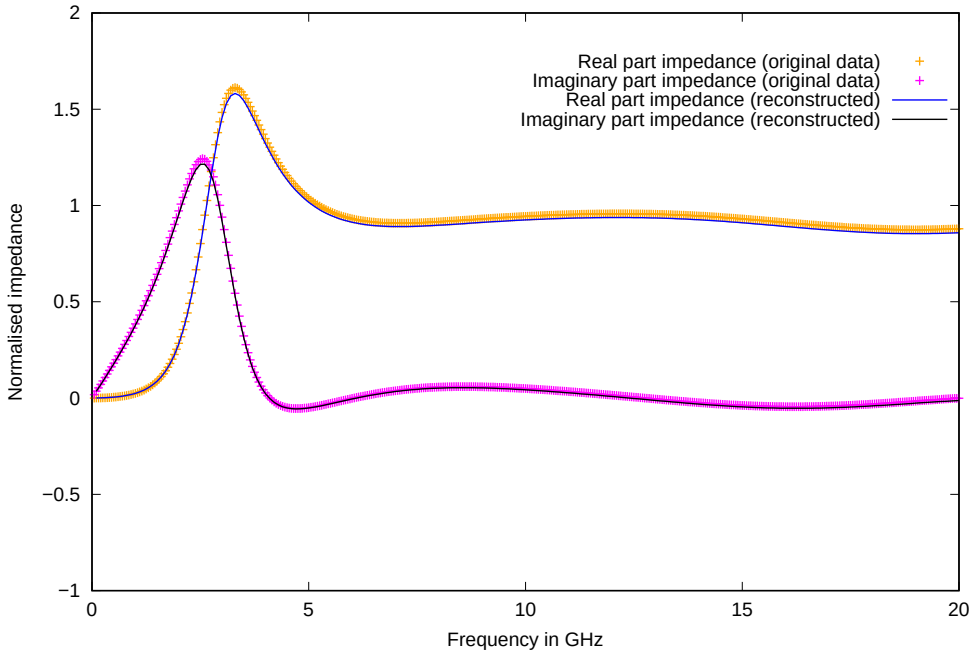

Figure 3-66: *Example 7, Gustavsen variant 1, for $N = 18$. Impedance function fit with Z_{in} as a function of frequency in GHz.*

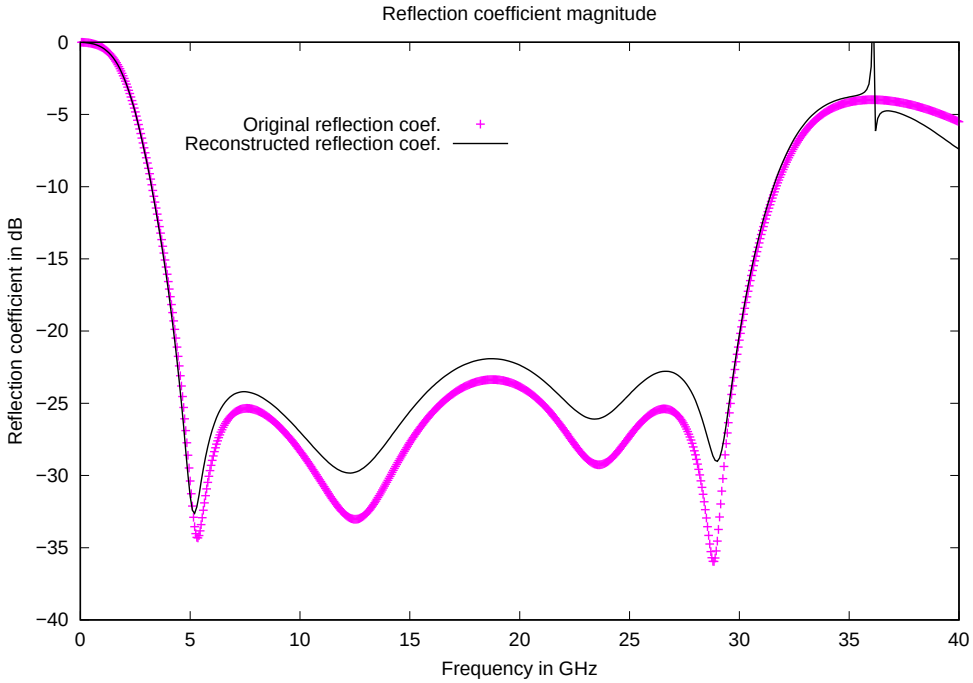

Figure 3-67: *Example 7, Gustavsen variant 1, for $N = 18$. Reflection coefficient magnitude in dB as a function of frequency in GHz.*

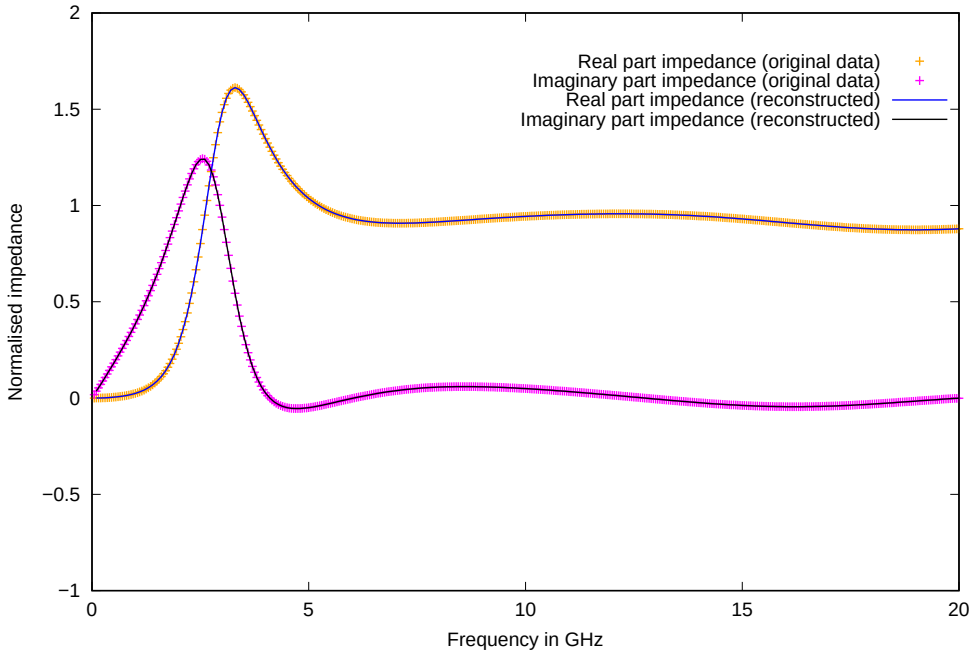

Figure 3-68: *Example 7, Null-space method, for $N = 18$. Impedance function fit with Z_{in} as a function of frequency in GHz.*

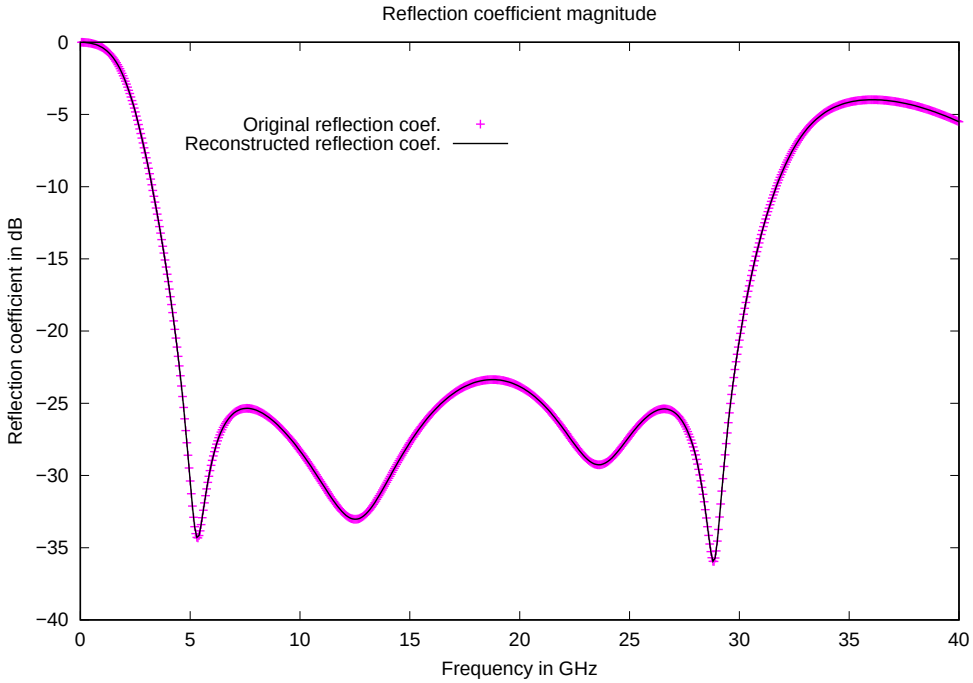

Figure 3-69: *Example 7, Null-space method, for $N = 18$. Reflection coefficient magnitude in dB as a function of frequency in GHz.*

3.10 Conclusions

A number of variants of Gustavsen and Semlyen's original vector fitting algorithm [1] are considered in this report. These include Gustavsen's second variant [2] and some modifications. In the examples presented, the second variant and its modifications are generally not as robust as the original.

Other variants are also considered. Of these, the one referred to as the null-space method is appealing. This method is demonstrated to work well when the other variants fail, in particular when the model order is overestimated. The method has a control parameter, L, which defines the size of the null space in the algorithm. This can be varied so as to maximise the accuracy of the algorithm, but it is frequently unnecessary to do so with the algorithm demonstrating robustness over a wide range of values of L.

In general it appears that Gustavsen's first variant and the null-space method complement each other well. Gustavsen's variant is best suited to problems requiring model order reduction and for transmission line problems where the model order number is similar to or smaller than optimal. On the other hand the null-space method works well when the model order number is similar to or larger than optimal. Using both allows a high degree of confidence in determining the pole locations of infinite-order impedance models over restricted bandwidths.

4 References

1 *B. Gustavsen, A. Semlyen "Rational approximation of frequency domain responses by vector fitting"*, IEEE Trans on Power Delivery, Vol. 14, No. 3, July 1999.

2 *B. Gustavsen "Improving the pole relocating properties of vector fitting"*, IEEE Trans on Power Delivery, Vol. 21, No. 3, July 2006.

3 *B. Gustavsen, A. Semlyen "Simulation on transmission line transients using vector fitting and modal decomposition"* IEEE Tran on Power Delivery, Vol. 13, No. 2, April 1998.

4 *M. Ahmadi "A network synthesis based approach for frequency dependent network equivalents"*, PhD thesis, University of Manitoba, 2021.

5 *A. C. S. Lima, B. Gustavsen, A. B. Fernandez "Inaccuracies in network realisation of rational models due to finite precision of RLC branches"* International Conference on Power Systems Transients (IPST'07), Lyon, France, June 4-7, 2007.

6 *W. Hendrickx, D. Deschrijver, T. Dhaene "Some remarks on the vector fitting iteration"*, Progress in Industrial Mathematics at ECMI 2004, pp 134-138 (available on-line from citeseerx.ist.psu.edu).

7 *A. Kazemzadeh, A. Karlsson* "Multilayered wideband absorbers for oblique angle of incidence" Lund University report 2010, TEAT 7188.

8 *K. N. Rozanov "Ultimate thickness to bandwidth ratio of radar absorbers"*, IEEE Transactions Antennas and Propagation, Vol. 48, No. 8, August 2000, pp1230-1234.

9 *C. Brewitt-Taylor "Progress on fundamental limits on radar-absorbing structures"*, DERA (QinetiQ) report DERA/S&P/RAD/TR000238/1.0, UK Defence Evaluation and Research Agency, Farnborough, Hampshire, GU14 OLX, UK. March 2000.

10 *C. Brewitt-Taylor "Fundamental limitation on the performance of chiral radar absorbing materials"* 8th International Conference on Electromagnetics of complex media, September 27-29, 2000.

11 *D. Deschrijver, B. Gustavsen, T. Dhaene "Advancements in iterative methods for rational approximation in the frequency domain"*, IEEE Trans Power Delivery, Vol. 22, No. 3, July 2007.

A Least squares minimisation of the residual error E_2

The residual error terms $\varepsilon_i^{(2)}$, defined in (2-14), may be written as the vector of terms

$$\boldsymbol{\varepsilon}^{(2)} = [\varepsilon_1^{(2)}, \varepsilon_2^{(2)}, \ldots, \varepsilon_{N_f}^{(2)}]^T \tag{a-1}$$

so that in vector-matrix notation,

$$\boldsymbol{\varepsilon}^{(2)} = \boldsymbol{f} + A_s \boldsymbol{x}'_s \tag{a-2}$$

where

$$\boldsymbol{x}'_s = [\boldsymbol{r}; d, e]^T \tag{a-3}$$

Now consider two cases:

Case 1

Suppose that $\boldsymbol{x}'_s$ is given up to a single unknown constant α,

$$\boldsymbol{x}'_s = \alpha \boldsymbol{x}_s \tag{a-4}$$

as defined in (2-42) of section 2.5, where α is a real unknown quantity.

The residual error $E_2 = \boldsymbol{\varepsilon}^{(2)\star}.\boldsymbol{\varepsilon}^{(2)}$ is minimised as a function of α when,

$$\frac{\partial}{\partial \alpha}\Big((\boldsymbol{f} + \alpha A_s \boldsymbol{x}_s)^\dagger (\boldsymbol{f} + \alpha A_s \boldsymbol{x}_s)\Big) = 0 \tag{a-5}$$

so that,

$$\boldsymbol{f}^\dagger A_s \boldsymbol{x}_s + (A_s \boldsymbol{x}_s)^\dagger \boldsymbol{f} = -2\alpha (A_s \boldsymbol{x}_s)^\dagger (A_s \boldsymbol{x}_s) \tag{a-6}$$

If the complex vector $\boldsymbol{c}$ is defined by,

$$\boldsymbol{c} = A_s \boldsymbol{x}_s \tag{a-7}$$

then

$$\alpha = \frac{-(\boldsymbol{f}^\dagger.\boldsymbol{c} + \boldsymbol{c}^\dagger.\boldsymbol{f})}{2\boldsymbol{c}^\dagger.\boldsymbol{c}} = \frac{-\Re(\boldsymbol{f}^\dagger.\boldsymbol{c})}{\boldsymbol{c}^\dagger.\boldsymbol{c}} \tag{a-8}$$

where the "." represents the conventional dot product between vectors.

Case 2

Suppose, here, that there are a set of column vectors $\boldsymbol{x}_l^{(s)}$, each of dimension N_2, for $l = 1$ to N_s rather than just a single vector as above. These are chosen as the sub-vectors of the set of eigenvectors $\hat{\boldsymbol{x}}_i$ of (2-33) associated with small eigenvalue magnitudes. Specifically, if $\hat{\boldsymbol{x}}_l$ are the eigenvectors then each $\hat{\boldsymbol{x}}_l$ may be written in terms of its sub-vectors as,

$$\hat{\boldsymbol{x}}_l = [\boldsymbol{x}_l^{(u)}; \boldsymbol{x}_l^{(v)}; x_l^{(\tilde{d})}, x_l^{(d)}, x_l^{(e)}]^T \tag{a-9}$$

where, up to an unknown scaling factor, $\boldsymbol{x}^{(u)}$ is a column vector representing candidates for the first N tilde residue terms for $\sigma(s)$, $\boldsymbol{x}^{(v)}$ the next N residue candidates for $p(s)$ and the remaining terms candidates for $\tilde{d}$, $x_l^{(d)}$ and $x_l^{(e)}$ respectively. Of immediate interest are the N_s sub-vectors,

$$\boldsymbol{x}_l^{(s)} = [\boldsymbol{x}_l^{(v)}; x_l^{(d)}, x_l^{(e)}]^T \tag{a-10}$$

Now we require that linear combination of $\boldsymbol{x}_i^{(s)}$ that minimises E_2. In this case define the real rectangular matrix, of dimension $N_2 \times N_s$,

$$X_r = [\boldsymbol{x}_1^{(s)}, \boldsymbol{x}_2^{(s)}, \ ..., \boldsymbol{x}_{N_s}^{(s)}] \tag{a-11}$$

and the $N_f \times N_s$ complex matrix,

$$C = A_s X_r \tag{a-12}$$

The required vector $\boldsymbol{x}_s'$ that minimises E_2 is given by,

$$\boldsymbol{x}_s' = X_r \boldsymbol{\alpha} \tag{a-13}$$

where

$$\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ ..., \alpha_{N_s}]^T \tag{a-14}$$

represents the real vector of coefficients α_i associated with each $\boldsymbol{x}_i^{(s)}$. The objective is then to minimise E_2 with respect to each of these coefficients in order to obtain an expression for $\boldsymbol{\alpha}$. The error,

$$|\boldsymbol{\varepsilon^{(2)}}|^2 = \boldsymbol{f}^\dagger \boldsymbol{f} + \boldsymbol{\alpha}^T C^\dagger \boldsymbol{f} + \boldsymbol{f}^\dagger C \boldsymbol{\alpha} + \boldsymbol{\alpha}^T C^\dagger C \boldsymbol{\alpha} \tag{a-15}$$

but solution to the minimisation problem is best performed using index notation. The Einstein summation convention over repeated indices is adopted except with respect to the special index p where no summation is implied. So,

$$|\boldsymbol{\varepsilon^{(2)}}|^2 = f_i^\star f_i + \alpha_j C_{ij}^\star f_i + f_i^\star C_{ij} \alpha_j + \alpha_j C_{ij}^\star C_{ik} \alpha_k \tag{a-16}$$

Setting the partial derivatives $\partial |\boldsymbol{\varepsilon^{(2)}}|^2 / \partial \alpha_p = 0$ implies,

$$C_{ip}^\star f_i + f_i^\star C_{ip} + C_{ip}^\star C_{ik} \alpha_k|_{k \neq p} + \alpha_j C_{ij}^\star C_{ip}|_{j \neq p} + 2\alpha_p C_{ip}^\star C_{ip} = 0 \tag{a-17}$$

simplifying to,

$$C_{ip}^\star f_i + f_i^\star C_{ip} + C_{ip}^\star C_{ik} \alpha_k + C_{ip} C_{ij}^\star \alpha_j \tag{a-18}$$

Converting back to matrix notation,

$$-(C^\dagger \boldsymbol{f} + (C^\dagger \boldsymbol{f})^\star) = (C^\dagger C + (C^\dagger C)^\star) \boldsymbol{\alpha} \tag{a-19}$$

with the solution,

$$\boldsymbol{\alpha} = -[\Re(C^\dagger C)]^{-1} \Re(C^\dagger \boldsymbol{f}) \tag{a-20}$$

which is the generalisation of (a-8). Note that $C^\dagger C$ is complex Hermitian so $\Re(C^\dagger C) \neq C^\dagger C$ in general, but $\Re(C^\dagger C)$ is real symmetric. Its inverse is not guaranteed to exist, but its dimension N_s is taken to be much smaller than N_2 and the vectors comprising X_r are orthogonal when taken as eigenvectors of (2-33) so this is generally a safe assumption.

Once $\boldsymbol{\alpha}$ is determined, the residues $\boldsymbol{r}$ with d and e are determined from (a-13). The coefficients α_i select on the full eigenvectors in (a-9) and not just on the sub-vectors used to minimise E_2 so the remaining vector of residues $\tilde{\boldsymbol{r}}$ and the term $\tilde{d}$ are given by,

$$\tilde{\boldsymbol{r}} = \sum_{l=1}^{N_s} \alpha_l \boldsymbol{x}_l^{(u)} \tag{a-21}$$

and

$$\tilde{d} = \sum_{l=1}^{N_s} \alpha_l x_l^{(\tilde{d})} \tag{a-22}$$

The expressions (a-21) and (a-22), together with the previous estimates for q_i, are used to compute the new pole positions. Unlike the single null-eigenvalue method, the requirement for more than one α_l coefficient requires these updates.

Note that the dimension of X_r is $N_2 \times N_s$. If $\boldsymbol{\alpha}$ is to be uniquely defined it is necessary that $N_s \leq N_2$. So, any selection of N_s by thresholding must be constrained by this requirement. Also, N_s is defined on the first iteration and remains fixed throughout the iteration cycle. Although is is possible to vary N_s, depending on the changing spectrum of the matrix B from one iteration to the next, experiments indicate that this can lead to poor convergence. The reasons for this remain unclear.

B Alternative proof of Gustavsen's matrix method for the zeros of $\sigma(s)$

In [1] and [2] a companion matrix is constructed given the poles, residues and scaling factor $\tilde{d}$. The zeros of the rational function with these parameters are then determined as the eigenvalues of the matrix. The result is very useful and was derived from state equation theory in [3]. In this appendix a more direct proof is given, using the Weinstein-Aronszajn identity.

Let us suppose the real and complex poles and residues of the function $\sigma(s)$ are those previously defined in section 2.5. There are P pairs of conjugate poles with values,

$$q_i = q'_i + jq''_i \text{ , for } i = 1, P \tag{b-1}$$

with the remaining R poles real valued. Similarly for the residues,

$$\tilde{r}_i = \tilde{r}'_i + j\tilde{r}''_i \text{ , for } i = 1, P \tag{b-2}$$

with the remaining R residues real valued. The proposition is that the zeros of $\sigma(s)$, defined by

$$\sigma(s) = \sum_{m=1}^{N} \frac{\tilde{r}_m}{s - q_m} + \tilde{d} \tag{b-3}$$

are the eigenvalues of the real non-symmetric companion matrix H, of dimension $N \times N$, defined by

$$H = H^{(1)} - \frac{1}{\tilde{d}} H^{(2)} \tag{b-4}$$

where

$$H^{(1)} = \begin{pmatrix} \begin{bmatrix} q'_1 & q''_1 \\ -q''_1 & q'_1 \end{bmatrix} & & & & \\ & \begin{bmatrix} q'_2 & q''_2 \\ -q''_2 & q'_2 \end{bmatrix} & & 0 & \\ & & \ddots & & \\ & & & \begin{bmatrix} q'_P & q''_P \\ -q''_P & q'_P \end{bmatrix} & \\ & 0 & & & q_{2P+1} \\ & & & & & \ddots \\ & & & & & & q_{2P+R} \end{pmatrix} \tag{b-5}$$

and $H^{(2)}$ is the tensor product of the two vectors, $(\boldsymbol{h}_b)^T$ and $\boldsymbol{h}_c$,

$$H^{(2)} = (\boldsymbol{h}_b)^T \boldsymbol{h}_c \tag{b-6}$$

where,

$$\boldsymbol{h}_b = [(2,0), (2,0), \ldots, (2,0); 1, 1, \ldots, 1] \tag{b-7}$$

where there are P pairs of terms $(2,0)$ and R remaining 1 entries, and

$$\boldsymbol{h}_c = [(\tilde{r}'_1, \tilde{r}''_1), (\tilde{r}'_2, \tilde{r}''_2), \ ..., (\tilde{r}'_P, \tilde{r}''_P); \tilde{r}_{2P+1}, \tilde{r}_{2P+2}, \ ..., \tilde{r}_{2P+R}] \tag{b-8}$$

with all terms real.

To prove this result, first note that the zeros of (b-3) are the zeros of;

$$\sigma(s)/\tilde{d} = \sum_{m=1}^{N} \frac{R_m}{s - q_m} + 1 \tag{b-9}$$

where $R_m = \tilde{r}_m/\tilde{d}$ so it is only necessary to consider the case $\tilde{d} = 1$ and then re-scale. The eigenvalues, $\lambda = \lambda_i$, of H are determined by the characteristic equation,

$$det(H - \lambda I_N) = 0 \tag{b-10}$$

where I_N is the unit matrix of dimension $N = 2P + R$.

The Weinstein-Aronszajn identity states that A is a matrix of size $m \times n$ and B is a matrix of size $n \times m$ and provided that the product AB is of trace class, then

$$det(I_m + AB) = det(I_n + BA) \tag{b-11}$$

To begin, assume $\tilde{d} = 1$ so that

$$det(H - \lambda I_N) = det(H^{(1)} - \lambda I_N - (\boldsymbol{h}_b)^T \boldsymbol{h}_c) \tag{b-12}$$

where $(\boldsymbol{h}_b)^T$ is of dimension $N \times 1$ and $\boldsymbol{h}_c$ is of dimension $1 \times N$. Thus,

$$det(H - \lambda I_N) = det(H^{(1)} - \lambda I_N)\, det(I_N - (H^{(1)} - \lambda I_N)^{-1} (\boldsymbol{h}_b)^T \boldsymbol{h}_c) \tag{b-13}$$

since the matrix inverse $(H^{(1)} - \lambda I_N)^{-1}$ exists provided the zeros do not coincide with the poles q_i, which by assumption they do not. The Weinstein-Aronszajn identity is then applied to the second determinant, so that

$$det(H - \lambda I_N) = det(H^{(1)} - \lambda I_N)\, det(I_1 - \boldsymbol{h}_c (H^{(1)} - \lambda I_N)^{-1} (\boldsymbol{h}_b)^T) \tag{b-14}$$

where $I_1 = 1$ is the identity element of dimension one.

Since $H^{(1)} - \lambda I_N$ is of block diagonal form, its inverse is of the same block diagonal form,

$$(H^{(1)} - \lambda I_N)^{-1} = \begin{pmatrix} \begin{bmatrix} q'_1 - \lambda & q''_1 \\ -q''_1 & q'_1 - \lambda \end{bmatrix}^{-1} & & & & \\ & \begin{bmatrix} q'_2 - \lambda & q''_2 \\ -q''_2 & q'_2 - \lambda \end{bmatrix}^{-1} & & 0 & \\ & & \ddots & & \\ & & & \begin{bmatrix} q'_P - \lambda & q''_P \\ -q''_P & q'_P - \lambda \end{bmatrix}^{-1} & \\ & 0 & & & 1/(q_{2P+1} - \lambda) \\ & & & & & \ddots \\ & & & & & & 1/(q_{2P+R} - \lambda) \end{pmatrix}$$

and using the definitions for the complex q_i and complex $\tilde{r}_i$,

$$\boldsymbol{h}_c(H^{(1)} - \lambda I_N)^{-1}(\boldsymbol{h}_b)^T = \sum_{i=1}^{N} \frac{\tilde{r}_i}{q_i - \lambda} \tag{b-15}$$

The determinant $det(H^{(1)} - \lambda I_N)$ is given as the product of the determinant of each block diagonal entry so that,

$$det(H^{(1)} - \lambda I_N) = \prod_{i=1}^{P} \left((q_i' - \lambda)^2 + (q_i'')^2\right) \prod_{i=2P+1}^{R} (q_i - \lambda) = \prod_{i=1}^{N} (q_i - \lambda) \tag{b-16}$$

hence

$$det(H - \lambda I_N) = \prod_{i=1}^{N} (q_i - \lambda) \,.\, \left(1 + \sum_{i=1}^{N} \frac{\tilde{r}_i}{\lambda - q_i}\right) \tag{b-17}$$

Setting this to zero, as the characteristic equation, implies that,

$$1 + \sum_{i=1}^{N} \frac{\tilde{r}_i}{\lambda - q_i} = 0 \tag{b-18}$$

so that the values $\lambda = \lambda_i$ which solve this equation are exactly those for which solve $\sigma(\lambda) = 0$, i.e. the zeros of $\sigma(s)$. When $\tilde{d} \neq 1$, rescaling $\tilde{r}_i \to \tilde{r}_i/\tilde{d}$ in (b-18) induces (b-4) and *vice-versa* so the general proposition is proved. QED.