



Recursive versus nonrecursive Richardson algorithms: systematic overview, unified frameworks and application to electric grid power quality

Downloaded from: <https://research.chalmers.se>, 2025-12-05 03:46 UTC

Citation for the original published paper (version of record):

Stotsky, A. (2022). Recursive versus nonrecursive Richardson algorithms: systematic overview, unified frameworks and application to electric grid power quality monitoring. *Automatika*, 63(2): 328-337.
<http://dx.doi.org/10.1080/00051144.2022.2039989>

N.B. When citing this work, cite the original published paper.



Recursive versus nonrecursive Richardson algorithms: systematic overview, unified frameworks and application to electric grid power quality monitoring

Alexander Stotsky

Department of Computer Science and Engineering, Chalmers University and University of Gothenburg, Gothenburg, Sweden

ABSTRACT

Sufficiently accurate, fast and computationally efficient solution of the system of linear equations is required in many estimation problems. Richardson iteration is one of the main solvers for linear equations, which provides optimization possibilities for time critical and accuracy critical applications. Convergence rate improvement and reduction of the computational complexity of the Richardson iteration are the most important problems in the area. The introduction of Newton–Schulz iterations is the efficient way for convergence rate improvement and the paper starts with systematic overview of the high-order Newton–Schulz matrix inversion algorithms. In addition, the unified framework for recursive computationally efficient convergence accelerators and error models for a number of combinations of Richardson and Newton–Schulz iterations is developed. A new nonrecursive parameter estimation concept is introduced and compared in this paper with recursive estimation. Recursive and nonrecursive Richardson algorithms together with the standard LU decomposition method were applied to the electric grid power quality monitoring problem. The algorithms were tested for the detection of the sag and swell signatures in the voltage and current signals on real data in three-phase power system. Nonrecursive Richardson algorithms which save close to half of the computational time compared to LU decomposition method were recommended for power quality monitoring applications.

ARTICLE HISTORY

Received 29 July 2021
Accepted 4 February 2022

KEYWORDS

Richardson algorithms;
Newton–Schulz matrix
inversion algorithms; power
quality monitoring

1. Introduction

1.1. Restricted linear system of equations and Richardson iteration

The solution of the system of linear equations

$$A\theta_* = b \quad (1)$$

with respect to θ_* with SPD (Symmetric and Positive Definite) (and ill-conditioned in many cases) matrix A is required in many application areas such as control, system identification, signal processing, statistics as well as in many big data applications, machine learning and in many others, [1–4].

The iterative Richardson method for solving (1) is often preferable due to the *optimization possibility* (achieved via a proper choice of the number of iterations), robustness, less processor time and memory space compared to direct methods. The trade-off between parameter estimation accuracy and computational burden allows different optimization of Richardson algorithms for approximate computing in time and accuracy critical applications. Convergence rate improvement and reduction of the computational complexity of the Richardson iteration are the most important problems in the area.

The convergence rate of the Richardson iteration can be improved via the introduction of fast iterative matrix inversion algorithms, [5–8] such as high order Newton–Schulz algorithms, [9–19]. Iterative algorithms for calculations of the inverses and different generalized inverses are widely used in cryptography, control systems, signal and image processing and in many other areas. Notice that practical applications of the iterative matrix inversion methods are hampered by the lack of general unified description. Mini-tutorial with general unified description of the iterative matrix inversion algorithms with convergence analysis is presented in this paper.

Convergence rate improvement is usually associated with additional computational burden. In order to reduce the computational complexity the following algorithms should be developed: (a) the recursive procedures associated with combinations of Newton–Schulz and Richardson iterations and (b) nonrecursive forms of the Richardson algorithm.

1.2. Nonrecursive Richardson algorithms

The Richardson iteration can be presented in a nonrecursive form, where the estimate is the function of

initial estimate and initial power series and all the quantities are calculated once only, [5]. The Richardson algorithm written in a nonrecursive form can be implemented on parallel computational units using sequential matrix vector multiplications only. Accuracy critical applications require a relatively large number of iterations of the recursive algorithm for the ill-conditioned case. This results in a large number of matrix vector multiplications in a nonrecursive counterpart. The approach is computationally efficient for large-scale systems and less efficient for small and medium size matrices. The algorithm should be selected according to the properties of the system in each application.

Factorizations of the power series, see for example [8,15,20] in combination with matrix vector multiplications can be applied for the improvement of the computational efficiency of the solution taking into account the size of the matrix A .

1.3. Application: electric grid power quality monitoring

Power quality disturbances such as sag and swell events may cause significant damages and problems to critical equipment connected to the network. Fast detection with possible mitigation of these events is one of the most important problems in the area, [21–24]. A number of approaches such as peak voltage detection, root mean square method, Fourier transformation and many others have been developed for sag and swell detection. Voltage peak detection is not robust in the presence of harmonics and the detection methods, which use the windowing technique can be slow for a sufficiently large window size, [25]. The approach presented in this paper allows to use short enough windows for the estimation of the frequency contents of oscillating signals and rapid detection of the disturbances. The estimation involves solution of (1) with positive definite and ill-conditioned (due to a small window size) information matrix A by using Richardson algorithms with improved convergence rate.

The algorithms developed in this paper have been tested for the detection of the sag and swell signatures of real failure event in voltage and current signals in three-phase power systems, [26]. Nonrecursive Richardson algorithms which can be optimized via a proper choice of the number of steps for approximate computing save close to half of the computational time compared to the LU decomposition method which does not allow optimization of the performance. Detection performance improvement associated with computational savings of the nonrecursive Richardson method is an essential contribution to time critical power quality monitoring applications.

1.4. Structure of the paper and main contributions

This paper starts with short systematic overview (presented as a mini tutorial) of iterative matrix inversion algorithms with unified convergence analysis in Section 2. In addition to unified description of known algorithms, a new high order iteration with memory based convergence accelerator is introduced. A new unified framework for recursive computationally efficient accelerators for the Richardson algorithm presented in the form of Newton–Schulz iterations, see Section 3 is the first main contribution of the paper. New nonrecursive estimators (the second main contribution of the paper) are introduced in Section 4 and compared to recursive estimators in Section 5. Grid power quality monitoring application is considered in Section 6, where nonrecursive and recursive estimators and the standard LU decomposition method are compared. The paper ends with brief conclusions and future outlook in Section 7.

2. Overview of iterative matrix inversion algorithms: mini tutorial

A brief overview of iterative matrix inversion algorithms with convergence analysis is presented in this Section. The overview is presented in the systematic way (from simple algorithms to more sophisticated ones) in the chronological sequence with convergence analysis and starts with the second order Newton–Schulz algorithms in Section 2.1 followed by the description of high order Newton–Schulz algorithms in Section 2.2. Computationally efficient Durand iterations are described in Section 2.3. High-order Newton–Schulz algorithms with improved convergence rate are presented in Section 2.4, where the combination of high order Newton–Schulz and Durand algorithms is also discussed. Finally, a new high-order Newton–Schulz iteration with memory based convergence accelerator is presented in Section 2.4.4.

Iterative algorithms for calculations of the inverses and different generalized inverses are widely used in cryptography, control systems, signal and image processing and in many other areas.

2.1. Second order Newton–Schulz iterations

Consider the following iterations, [9–12]:

$$G_k = G_{k-1} + F_{k-1}G_{k-1}, \quad F_k = I - G_k A \quad (2)$$

where G_k is the estimate of the inverse of the matrix A and I is the identity matrix, $k = 1, 2, \dots$. The evaluation of the error in the first step F_1 yields:

$$\begin{aligned} F_1 &= I - (G_0 + F_0 G_0) A = I - (I + F_0) G_0 A \\ &= I - (I + F_0) (I - F_0) = I - I + F_0^2 = F_0^2 \end{aligned}$$

$$F_2 = F_1^2 = F_0^4$$

which gives the error model:

$$F_k = F_0^{2^k} \quad (3)$$

Algorithm (2) which is also known as the Newton–Schulz–Hotelling matrix iteration, [10] estimates the inverse of the matrix A via minimization of the estimation error F_k provided that the spectral radius of initial matrix

$$\rho(F_0) = \rho(I - G_0 A) < 1 \quad (4)$$

is less than one, where G_0 is the preconditioner (initial guess of A^{-1}). The preconditioner has especially a simple form for SPD matrices which can be split as $A = S - D$. The spectral radius of the matrix $I - S^{-1}A$ is less than one, $\rho(I - S^{-1}A) = \rho(S^{-1}D) < 1$ provided that the preconditioner is chosen as $G_0 = S^{-1} = I/\alpha$ with $\alpha = \|A\|_\infty/2 + \varepsilon$, where $\|\cdot\|_\infty$ is the maximum row sum matrix norm, $\varepsilon > 0$, [3].

The proof of convergence is presented in [9–12], where the algorithm (2) is written in the following form: $G_k = G_{k-1} + G_{k-1}F_{k-1}$, where $F_k = I - AG_k$.

The order of the iterative algorithm can be increased for convergence rate improvement by the introduction of the power series in the algorithm (2).

2.2. High-order Newton–Schulz iterations

High-order algorithms (which are also known as hyper-power iterative algorithms, which are widely used for calculation of generalized inverses) can be presented in the following form for $n = 2, 3, \dots$, [13–15]:

$$\begin{aligned} G_k &= \sum_{j=0}^{n-1} F_{k-1}^j G_{k-1} \\ &= (I + F_{k-1} + F_{k-1}^2 + \dots + F_{k-1}^{n-1})G_{k-1} \end{aligned} \quad (5)$$

or in the form presented in [6,14]:

$$G_k = \sum_{j=0}^{n-1} F_{k-1}^j G_{k-1} = G_{k-1} + F_{k-1} \sum_{j=0}^{n-2} F_{k-1}^j G_{k-1} \quad (6)$$

The error F_k in (5) can be written in the following form:

$$\begin{aligned} F_k &= I - G_k A \\ &= I - (I + F_{k-1} + F_{k-1}^2 + \dots + F_{k-1}^{n-1})G_{k-1}A \\ &= I - (I + F_{k-1} + F_{k-1}^2 + \dots + F_{k-1}^{n-1})(I - F_{k-1}) \\ &= F_{k-1}^n \\ F_k &= F_0^{n^k} \end{aligned} \quad (7)$$

Notice that algorithm (5) is written in following form in [13–15] (and it is widely used in this form):

$$\begin{aligned} G_k &= G_{k-1} \sum_{j=0}^{n-1} F_{k-1}^j \\ &= G_{k-1}(I + F_{k-1} + F_{k-1}^2 + \dots + F_{k-1}^{n-1}) \end{aligned} \quad (8)$$

with $F_k = I - AG_k$. Notice that the error $F_k = I - G_k A$ is better aligned with the Richardson iteration and therefore is considered further in this paper.

2.3. Durand iterations

The computationally efficient matrix inversion method described by Durand, [16,17] can be derived from the relation associated with the initial error in the Newton–Schulz approach, $A^{-1} = F_0 A^{-1} + G_0$ by the substitution of G_k instead of A^{-1} as follows:

$$G_k = F_0 G_{k-1} + G_0 \quad (9)$$

$$G_k - A^{-1} = F_0(G_{k-1} - A^{-1}) \quad (10)$$

$$F_k = I - G_k A = F_0^{k+1} \quad (11)$$

The error models (10), (11) are valid for the algorithm (9). Higher order Durand iterations are discussed in [4,17].

Notice that the convergence rate of Newton–Schulz iterations is significantly higher than the convergence rate of Durand iterations. However, Durand iteration (9) requires only one matrix product in each step.

Notice also that the convergence performance can be enhanced by the introduction of the initial power series as preconditioning G_0 (which is calculated only once) in Newton–Schulz and Durand iterations [4,18]. The order of initial power series can be chosen sufficiently high, which essentially improves the convergence rate. Convergence rate can also be improved via combination of two Newton–Schulz iterations, which also cover Durand iterations as special case, see next Section 2.4.

2.4. Convergence rate improvement of high order Newton–Schulz iterations

2.4.1. Iterative method with improved convergence rate

The improved iterative method can be written in the following form, [19]:

$$Z_k = \sum_{j=0}^{n-1} (I - Z_{k-1}A)^j Z_{k-1} \quad (12)$$

$$G_k = G_{k-1} + (I - G_{k-1}A)Z_k \quad (13)$$

where both Z_k and G_k are the estimates of the matrix inverse and $n = 2, 3, \dots$ is the order. Multiplication

of (13) by A together with evaluation of the power series (12) yields:

$$\underbrace{I - G_k A}_{F_k} = \underbrace{I - G_{k-1} A}_{F_{k-1}} - \underbrace{(I - G_{k-1} A)}_{F_{k-1}} \underbrace{Z_k A}_{I - L_{k-1}^n}$$

and algorithm (12) and (13) has the following error model, [8]:

$$L_k = L_{k-1}^n, \quad L_k = I - Z_k A \quad (14)$$

$$F_k = F_{k-1} L_{k-1}^n, \quad F_k = I - G_k A \quad (15)$$

Introduction of the power series for F_{k-1} improves further the convergence rate.

2.4.2. Combination of two high order Newton-Schulz iterations

Two Newton-Schulz iterations can be combined as follows, [8]:

$$\Gamma_k = I - L_{k-1} A \quad (16)$$

$$L_k = \left\{ \sum_{j=0}^{n-1} \Gamma_k^j \right\} L_{k-1} \quad (17)$$

$$\Gamma_k^n = I - L_k A \quad (18)$$

$$G_k = \underbrace{L_k}_{\text{Newton-Schulz Iteration}} + \underbrace{\Gamma_k^n}_{\text{High-Order Convergence Accelerator}} \underbrace{\left\{ \sum_{j=0}^{n-1} F_{k-1}^j \right\}}_{\text{Newton-Schulz Iteration}} G_{k-1} \quad (19)$$

$$F_k = I - G_k A = \Gamma_k^n F_{k-1}^n \quad (20)$$

$$F_k = F_0^{k n^{k+1} + n^k}, \quad \text{for } n > 1 \quad (21)$$

where $L_0 = \sum_{j=0}^{n-1} \Gamma_0^j S^{-1}$, $\Gamma_0 = F_0 = I - S^{-1} A$, where $G_0 = S^{-1}$ satisfies (4). Notice that Durand iteration (9) can be derived from (16)–(19) with $n = 1$. The error model (24) can be evaluated as follows:

$$\begin{aligned} F_k = I - G_k A &= \underbrace{I - L_k A}_{\Gamma_k^n} - \Gamma_k^n \sum_{j=0}^{n-1} F_{k-1}^j \underbrace{G_{k-1} A}_{I - F_{k-1}} \\ &= \Gamma_k^n F_{k-1}^n \end{aligned} \quad (22)$$

The error model (21) is obtained via explicit evaluation of (22). The algorithm (16)–(21) has a faster convergence due to the high order error F_{k-1}^n in the error model (20) compared to algorithm (12) and (13) which has the error model (15) with the first order error F_{k-1} .

The algorithm (16)–(19) has two independent computational parts, where the first part is associated with the calculations of Γ_k , L_k and Γ_k^n , where $\Gamma_k = \Gamma_{k-1}^n$, and the term $\sum_{j=0}^{n-1} F_{k-1}^j G_{k-1}$ is calculated in the second part. This algorithm is suitable for parallel implementation.

2.4.3. Nested Newton-Schulz iterations

The following new algorithm with nested Newton-Schulz loops and improved convergence rate extends the framework described above as follows:

$$\Gamma_k = I - L_{k-1} A \quad (23)$$

$$r \text{ loops } \begin{cases} L_k = \left\{ \sum_{j=0}^{n-1} \Gamma_k^j \right\} L_{k-1}, \quad \Gamma_k^n = I - L_k A \\ P_k = \left\{ \sum_{j=0}^{n-1} \Gamma_k^{nj} \right\} L_k, \quad \Gamma_k^{n^2} = I - P_k A \\ \dots \end{cases}$$

$$\begin{aligned} G_k &= \underbrace{T_c}_{\text{Composite Newton-Schulz Iteration}} \\ &+ \underbrace{\Gamma_c}_{\text{High-Order Convergence Accelerator}} \underbrace{\left\{ \sum_{j=0}^{n-1} F_{k-1}^j \right\}}_{\text{Newton-Schulz Iteration}} G_{k-1} \end{aligned} \quad (24)$$

$$T_c = L_k + \Gamma_k^n P_k + \dots,$$

$$\Gamma_c = I - T_c A = \Gamma_k^{\frac{(nr+1-n)}{(n-1)}} \quad (25)$$

$$F_k = I - G_k A = \Gamma_c F_{k-1}^n \quad (26)$$

$$F_k = F_0^{kn^{k+1} \frac{(nr-1)}{(n-1)} + n^k} \quad \text{for } n > 1 \quad (27)$$

where $L_0 = \sum_{j=0}^{n-1} \Gamma_0^j S^{-1}$, $\Gamma_0 = F_0 = I - S^{-1} A$, $G_0 = S^{-1}$ and $r = 1, 2, 3, \dots$ is the number of Newton-Schulz loops.

2.4.4. High order Newton-Schulz iteration with memory based convergence accelerator

The high order iterative method with memory based convergence accelerator, where the estimate in the step k is calculated using the estimates from the steps $k-1$ and $k-2$ and the error model accumulates multiplicatively inversion errors calculated in the previous steps can be written in the following form:

$$F_{k-1}^n = I - \sum_{j=0}^{n-1} F_{k-1}^j G_{k-1} A, \quad n = 1, 2, \dots \quad (28)$$

$$\begin{aligned} G_k &= G_{k-1} + P_{k-1} \left\{ \underbrace{\left(I + F_{k-1}^n \right) \sum_{j=0}^{n-1} F_{k-1}^j G_{k-1}}_{\sum_{j=0}^{2n-1} F_{k-1}^j G_{k-1}} \right. \\ &\quad \left. - \sum_{j=0}^{n-1} F_{k-2}^j G_{k-2} \right\} \end{aligned} \quad (29)$$

$$P_k = P_{k-1} F_{k-1}^n, \quad (30)$$

$$P_1 = F_0^n = I - \sum_{j=0}^{n-1} F_0^j G_0 A$$

$$F_k = I - G_k A = P_k F_{k-1}^n = \underbrace{F_0^n F_1^n \cdots F_{k-2}^n}_{\text{Memory Based Convergence Accelerator}} F_{k-1}^{2n} \quad (31)$$

where $G_1 = (I + F_0^n) \sum_{j=0}^{n-1} F_0^j G_0$ and G_0 satisfies (4). The error model (31) is obtained by multiplication of (29) by A which results in the error model $F_k = F_{k-1} + P_{k-1} \{F_{k-1}^{2n} - F_{k-2}^n\}$ with subsequent evaluation of P_{k-1} using (30). The algorithm (28)–(31) is new high-order method with a memory based convergence accelerator, which has especially simple form for $n = 1$, and a similar algorithm with memory can be found in [27].

Interestingly enough that the algorithm (28)–(31) represents fast power series expansion. The evaluation of one step of the algorithm (29) for $k = 2$ by taking into account that $G_1 = \sum_{j=0}^{2n-1} F_0^j G_0$ and $F_1 = F_0^{2n}$ is presented as follows:

$$G_2 = \sum_{j=0}^{2n-1} F_0^j G_0 + F_0^n \left\{ \sum_{j=0}^{4n^2-1} F_0^j G_0 - \sum_{j=0}^{n-1} F_0^j G_0 \right\}$$

$$= \sum_{j=0}^{2n-1} F_0^j G_0 + \sum_{j=2n}^{4n^2+n-1} F_0^j G_0 = \sum_{j=0}^{4n^2+n-1} F_0^j G_0$$

$$F_2 = I - G_2 A = F_0^n F_1^{2n} = F_0^{4n^2+n} \quad (32)$$

and results in the error model (32) which verifies the model (31). The expansion can be continued for the next steps for complete verification of the error model.

The evaluation of the power series shows the idea of the algorithm. The power series $\sum_{j=0}^{n-1} F_{k-2}^j G_{k-2}$ is used for truncation of the series $(I + F_{k-1}^n) \sum_{j=0}^{n-1} F_{k-1}^j G_{k-1}$ so that the resulting series multiplied by P_{k-1} extends the power series coming from the previous step G_{k-1} .

3. Richardson iteration: recursive accelerators and error models

Consider the following general form, [5] of the Richardson iteration (33):

$$\theta_k = \theta_{k-1} - \overbrace{\left\{ \sum_{j=0}^{\omega_k-1} F_0^j \right\}}^{V_k} G_0 \{A\theta_{k-1} - b\},$$

$$F_0 = I - G_0 A \quad (33)$$

$$\tilde{\theta}_k = F_0^{\omega_k} \tilde{\theta}_{k-1}, \quad \tilde{\theta}_k = \theta_k - \theta_* \quad (34)$$

$$\tilde{\theta}_k = F_0^{\sum_{j=1}^k \omega_j} \tilde{\theta}_0, \quad \tilde{\theta}_0 = \theta_0 - \theta_* \quad (35)$$

and error models (34), (35), where θ_k is the estimate of unknown parameters θ_* defined in (1) and ω_k is associated with the expansion rate of the power series, and V_k is the convergence accelerator.

Specifying the parameters ω_k and $\sum_{j=1}^k \omega_j$ the family of estimation algorithms with error models (34) and (35) which allow individual and comparative performance assessment is obtained, see Table 1. The table represents the systematic view (implementation ready tool-kit) for convergence rate improvement of the Richardson iteration with Newton–Schulz type recursive computationally efficient accelerators.

The first item in the table is the combination of Richardson and Newton–Schulz iterations described in [6,7], whereas the second item is recursive realization of the combination described in [6]. The third and the fourth items are described in [4,8]. The table can be further extended by integration of the matrix inversion algorithms presented in Section 2 in the Richardson iteration.

Moreover, a stepwise evaluation of the algorithm with the highest convergence rate (the third item in the table), which shows the idea of fast power series expansion for the convergence accelerator and validation of the error model (34) is presented below.

3.1. Stepwise evaluation of the convergence accelerator V_k , (3) in Table 1

The evaluation of the convergence accelerator V_k , starts with the evaluation of the power series L_1 and G_1 for the first step $k = 1$:

$$L_1 = \left\{ \sum_{j=0}^{n-1} F_0^{nj} \right\} \left\{ \sum_{j=0}^{n-1} F_0^j \right\} S^{-1} = \left\{ \sum_{j=0}^{n^2-1} F_0^j \right\} S^{-1} \quad (36)$$

$$G_1 = \left\{ \sum_{j=0}^{n-1} F_0^{nj} \right\} \left[\left\{ \sum_{j=0}^{2n-1} F_0^j \right\} S^{-1} - \sum_{j=0}^{n-1} F_0^j S^{-1} \right]$$

$$+ \sum_{j=0}^{n-1} F_0^j S^{-1} = \left\{ \sum_{j=0}^{n^2+n-1} F_0^j \right\} S^{-1} \quad (37)$$

and the convergence accelerator V_k is evaluated as follows:

$$V_1 = L_1 + \Gamma_1^n \left\{ \sum_{j=0}^{n-1} F_1^j \right\} G_1 = \underbrace{\left\{ \sum_{j=0}^{n^2-1} F_0^j \right\}}_{L_1} S^{-1} \quad (38)$$

Table 1. Recursive convergence accelerators V_k for Richardson iteration and error models.

N	Recursive parameter estimation $\theta_k = \theta_{k-1} - V_k \{A\theta_{k-1} - b\}, \rho(I - S^{-1}A) < 1$
1	$\omega_k = n^k, \sum_{j=1}^k \omega_j = \frac{n^{k+1} - n}{n-1}, n = 2, 3, \dots$ $V_0 = S^{-1}, V_k = \sum_{j=0}^{n-1} F_{k-1}^j V_{k-1}, F_k = I - V_k A$
2	$\omega_k = n^{k+1}, \sum_{j=1}^k \omega_j = \frac{n(n^{k+1} - n)}{n-1}, n = 2, 3, \dots$ $F_0 = I - S^{-1}A, V_0 = \sum_{j=0}^{n-1} F_0^j S^{-1}$ $V_k = \sum_{j=0}^{n-1} (I - V_{k-1}A)^j V_{k-1}$
3	$\omega_k = k n^{k+2} + 2 n^{k+1}, n = 2, 3, \dots,$ $\sum_{j=1}^k \omega_j = n^2 \frac{k n^{k+2} - (k-1) n^{k+1} - 2 n^k - n + 2}{(n-1)^2}$ $L_0 = \sum_{j=0}^{n-1} \Gamma_0^j S^{-1}, \Gamma_0 = F_0 = I - S^{-1}A, G_0 = S^{-1}$ $V_0 = (2I - L_0 A) L_0 = \left\{ \sum_{j=0}^{2n-1} F_0^j \right\} S^{-1}$ $\Gamma_k = \Gamma_{k-1}^n, L_k = \left\{ \sum_{j=0}^{n-1} \Gamma_k^j \right\} L_{k-1}, \Gamma_k^n = I - L_k A$ $G_k = \left\{ \sum_{j=0}^{n-1} \Gamma_k^j \right\} \left[V_{k-1} - \sum_{j=0}^{n-1} F_{k-1}^j G_{k-1} \right] + \sum_{j=0}^{n-1} F_{k-1}^j G_{k-1}$ $F_k = I - G_k A, V_k = L_k + \Gamma_k^n \left\{ \sum_{j=0}^{n-1} F_k^j \right\} G_k$
4	$\omega_k = k + 2, \sum_{j=1}^k \omega_j = \frac{k^2 + 5k}{2}, n = 1$ $F_0 = I - S^{-1}A, V_0 = (I + F_0) S^{-1}, V_k = F_0 V_{k-1} + S^{-1}$

$$\begin{aligned}
& + \underbrace{\left\{ \sum_{j=0}^{n-1} F_0^{n^2} \right\}}_{\Gamma_1^n} \underbrace{\left\{ \sum_{j=0}^{n-1} F_0^{(n^2+n)j} \right\}}_{G_1} \underbrace{\left\{ \sum_{j=0}^{n^2+n-1} F_0^j \right\}}_{S^{-1}} \\
& = \left\{ \sum_{j=0}^{n^3+2n^2-1} F_0^j \right\} S^{-1} = \left\{ \sum_{j=0}^{\omega_k-1} F_0^j \right\} G_0 \quad (39) \\
& \omega_k = k n^{k+2} + 2 n^{k+1} = n^3 + 2 n^2, \quad \text{for } k = 1 \quad (40)
\end{aligned}$$

The expansion principle showed in Equations (38)–(40) is valid in all the subsequent steps $k = 2, 3, \dots$

Another way of parameter estimation is based on non-recursive form of the Richardson algorithm and factorization of the power series, described in the next section.

4. Nonrecursive Richardson estimators

Algorithm (33)–(35) can be presented in the following non-recursive form:

$$\theta_k = \theta_0 - \left\{ \sum_{j=0}^{\sum_{i=1}^k \omega_i - 1} F_0^j \right\} G_0 \{A\theta_0 - b\} \quad (41)$$

where the parameter vector θ_k is a function of the initial state θ_0 and power series $\{\sum_{j=0}^{\sum_{i=1}^k \omega_i - 1} F_0^j\} G_0$ only and all the quantities in (41) are calculated once only.

Algorithm (41) can be implemented using sequential matrix vector multiplications only, [5] with power series factorization of the polynomial $\sum_{j=0}^{\sum_{i=1}^k \omega_i - 1} F_0^j$ with respect to $(G_0 A)$ described for example in [8].

Accuracy critical applications require a relatively large number of iterations for the ill-conditioned case and in finite digit calculations. Accuracy requirements and requirements imposed on the convergence rate result in a high order of the polynomial $\sum_{j=0}^{\sum_{i=1}^k \omega_i - 1} F_0^j$, which involves a large number of matrix vector multiplications. Taking into account the size of the matrix A and the number of iterations to reach a certain accuracy a suitable method for factorization of the power series $\{\sum_{j=0}^{\sum_{i=1}^k \omega_i - 1} F_0^j\} G_0$ can be found in order to reduce the large number of matrix vector multiplications.

The procedure of the application of factorization based estimators can be described as follows:

- (1) The final step k_* is determined using the error model (35), pre-specified parameter estimation accuracy and the functions $\sum_{j=1}^k \omega_j$ presented in Table 1.
- (2) The power series $\{\sum_{j=0}^{\sum_{i=1}^{k_*} \omega_i - 1} F_0^j\} G_0$ which corresponds to the final step k_* (determined in the previous step) is factorized using the factorization tool-kit, [8] taking into account the size of matrix A and providing the best combination of matrix by matrix and matrix vector multiplications.
- (3) The final parameter estimate θ_{k_*} which guarantees pre-specified accuracy is calculated with (41) and the power series factorized in the previous step.

The procedure provides nonrecursive estimation of the parameter vector, which is more computationally efficient than recursive estimators for relatively small number of iterations.

5. Comparison of recursive and nonrecursive estimators

The recursive estimator presented as item 2 in Table 1 for the second order, $n = 2$ is chosen for comparison. The implementation of the algorithm requires 2 mmm

(matrix by matrix multiplications) in each step and matrix vector multiplications (mvm).

The nonrecursive counterpart has the form (41) with the following factorization $\{\sum_{j=0}^k \omega_i^{j-1} F_0^j\} G_0 = \{\sum_{j=0}^{2^{k+1}-3} F_0^{2j}\} \{I + F_0\} G_0$, [8] which requires two matrix products only and matrix vector multiplications.

The estimation of the parameter vector θ_* for the system (1) with ill-conditioned SPD information matrix A associated with the system with harmonic regressor and different number of frequencies is chosen for simulations. Several steps are considered for both recursive and nonrecursive algorithms providing exactly the same vector θ_k in both cases. The ratio of two computational times for the nonrecursive algorithm (in numerator) and the recursive algorithm (in denominator) multiplied by one hundred per cent is plotted as a function of the step number and the size of the matrix A (the number of frequencies for the system with harmonic regressor multiplied by two). Simulation results are presented in Figure 1. Nonrecursive algorithm is faster than the recursive one (time ratio is less than one hundred percent) for relatively small number of iterations.

Accuracy critical applications require a relatively large number of iterations for the ill-conditioned case. For the same convergence rate of both algorithms the large number of steps results in high-order polynomial $\sum_{j=0}^k \omega_i^{j-1} F_0^j$ in the nonrecursive algorithm. For example, for $k = 20$ and $n = 2$ the order is 4194300. The implementation of Richardson iteration for a relatively large number of steps requires very large number of matrix vector multiplications and the recursive algorithm is preferable in this case for small and medium size matrices, see Figure 1.

The advantage of the nonrecursive algorithm relative to the recursive algorithm is pronounced for the large-scale systems. However, nonrecursive estimators may have more significant error accumulation in finite digit calculations. In other words, the algorithms have distinct regions of applicability and can be applied for different cases. Further comparison is performed in the Section 6 for power system application.

6. Grid power quality monitoring

6.1. Estimation of the frequency content of the power system signals

Suppose that the measured signal of the power system (voltage or current) y_k can be presented in the following form:

$$y_k = \varphi_k^T \vartheta + \xi_k \quad (42)$$

where ϑ is the vector of unknown parameters and φ_k is the harmonic regressor presented in the following form:

$$\varphi_k^T = [\cos(q_0 k) \sin(q_0 k) \cos(2q_0 k) \sin(2q_0 k) \cdots \cos(mq_0 k) \sin(mq_0 k)] \quad (43)$$

where q_0 is the fundamental frequency of network (for example $q_0 = 50$ Hertz or $q_0 = 60$ Hertz), m is the number of harmonics, and ξ_k is a zero mean white Gaussian noise, $k = 1, 2, \dots$ is the step number.

The model of the signal (42) with adjustable parameters θ_{*k} is presented in the following form:

$$\hat{y}_k = \varphi_k^T \theta_{*k} \quad (44)$$

The signal y_k is approximated by the model \hat{y}_k in the least squares sense in each step k of moving window of a size s .

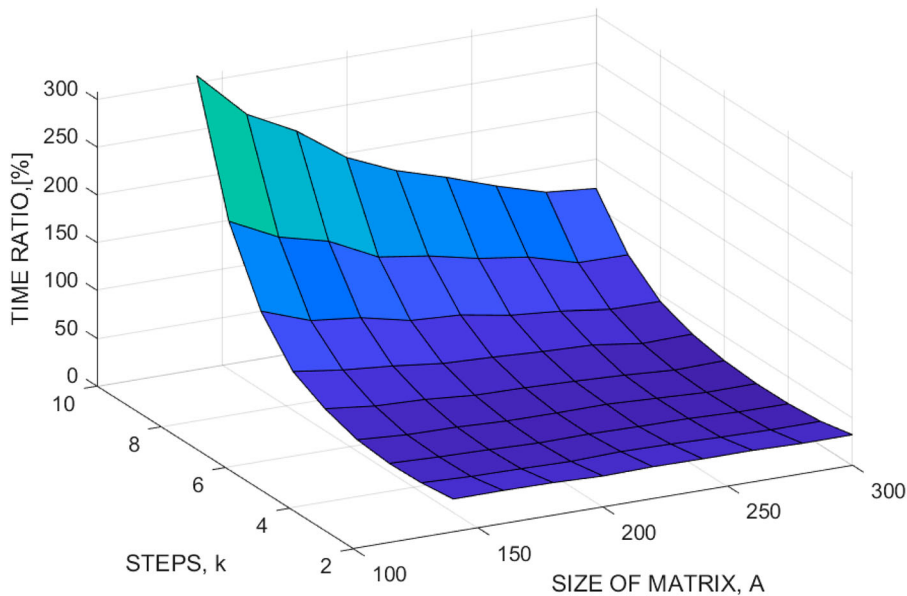


Figure 1. The steps $k = 2, \dots, 10$ are considered for recursive and nonrecursive algorithms described in Section 5 providing exactly the same vector θ_k in both cases. The ratio of two computational times for the nonrecursive algorithm (in numerator) and the recursive algorithm (in denominator) multiplied by 100% is plotted as a function of step number and the size of the matrix A (the number of frequencies for the system with harmonic regressor multiplied by two).

The estimation algorithm is based on minimization of the following error E_k ;

$$E_k = \sum_{p=k-(s-1)}^{p=k} (\hat{y}_p - y_p)^2 \quad (45)$$

for a fixed step k , where $k \geq s$.

The least squares solution for the estimation of the parameter vector θ_{*k} can be written as follows:

$$A_k \theta_{*k} = b_k \quad (46)$$

$$A_k = \sum_{p=k-(s-1)}^{p=k} \varphi_p \varphi_p^T \quad (47)$$

$$b_k = \sum_{p=k-(s-1)}^{p=k} \varphi_p y_p \quad (48)$$

where the matrix A_k is symmetric and positive-definite information matrix, and the parameter vector θ_{*k} , which satisfy (46) should be calculated in each step.

In power quality monitoring applications, the number of harmonics in voltage and current signals (higher harmonics are more pronounced in the current signals) is not large in three-phase systems and the size of the information matrix A_k is relatively small. Notice that the information matrix may change the size when new harmonics appear in the signals. Notice also that Equation (46) is exactly Equation (1) and the methods described above can be applied in this case.

6.2. Fault detection on real data

Power quality monitoring and detection of sag and swell events are time critical applications since the faults should be detected as soon as possible and appropriate action should be taken for power system protection. Monitoring is performed via the estimation of the frequency content of the voltage and current signals.

Real data obtained from the DOE/EPRI National Database of Power System Events, [26] are used for power quality monitoring. The event (Event ID 2827) which resulted in voltage and current sag and swell was considered. The event is associated with the fault on the lead cable caused by circuit breaker to trip on substation. Measured voltages and the currents which correspond to sag and swell in phases 'a', 'b' and 'c' are plotted with green, red and blue lines respectively in Figures 2(a,b).

It is assumed that the voltage and current signals can be described by the model (42) with harmonic regressor (43) and fundamental frequency of 60 Hz with four higher harmonics and unknown parameter vector ϑ .

The model of the signals is given by (44) with the vector of the parameters θ_{*k} which estimates unknown vector ϑ . The estimation problem is reduced to the

restricted linear system of Equation (46) which should be solved w.r.t. θ_{*k} in each step. The 10×10 matrix A_k is the SPD matrix in each step (which is verified by simulations). The preconditioner described in Section 2.1 can be easily calculated for this matrix in each step. Measured signals are well described by the model (42) with the small measurement noise ξ_k , which allows to choose short enough window ($s = 40$) for rapid detection. The choice of the window size represents the trade-off between the quality of the detection signal and detection time. Larger window sizes ($s \geq 54$) [22] imply that matrix A is strictly diagonally dominant for which better preconditioning, which reduces convergence time, is available, [6].

Amplitudes of the first harmonic of the measured signals recovered by the nonrecursive parameter estimation algorithm described in Section 5 are used for detection, see Figures 2(c,d). The figure shows that the sag and swell events are more pronounced on the current signal and detectable on both voltage and current signal.

Three algorithms (1) recursive estimator (2) nonrecursive estimator (both described in Section 5), and (3) the LU decomposition method are applied for parameter calculation of θ_{*k} in (46). The mean values of the computational times required for the estimation of the parameter vector are calculated for three algorithms. The computational time for the first two algorithms depends on the number of steps k_* and the minimal number of steps (four steps) which provides sufficient quality for the detection of the sag and swell events was chosen. Notice that the standard LU decomposition method (and parameter calculation method based on LU decomposition) does not have any parameters to optimize and provides worse parameter estimation accuracy than four steps of the first two methods. Simulation results show that the nonrecursive algorithm provides the best result and saves in average 59% of the computational time compared to the recursive algorithm and 43% compared to the standard LU decomposition method.

6.3. Discussion

The most significant computational burden associated with the implementation of the detection algorithm is associated with calculation of the parameters θ_{*k} in Equation (46) where the matrix A is a well-conditioned matrix for small enough window sizes. Therefore, the parameter vector θ_{*k} can be rapidly calculated using Richardson algorithms described above with a small number of steps. The nonrecursive algorithm provides the best computational performance. The result is in agreement with the results presented in Section 5, which show the advantages of the nonrecursive algorithm for a small number of steps and small size matrices, see Figure 1.

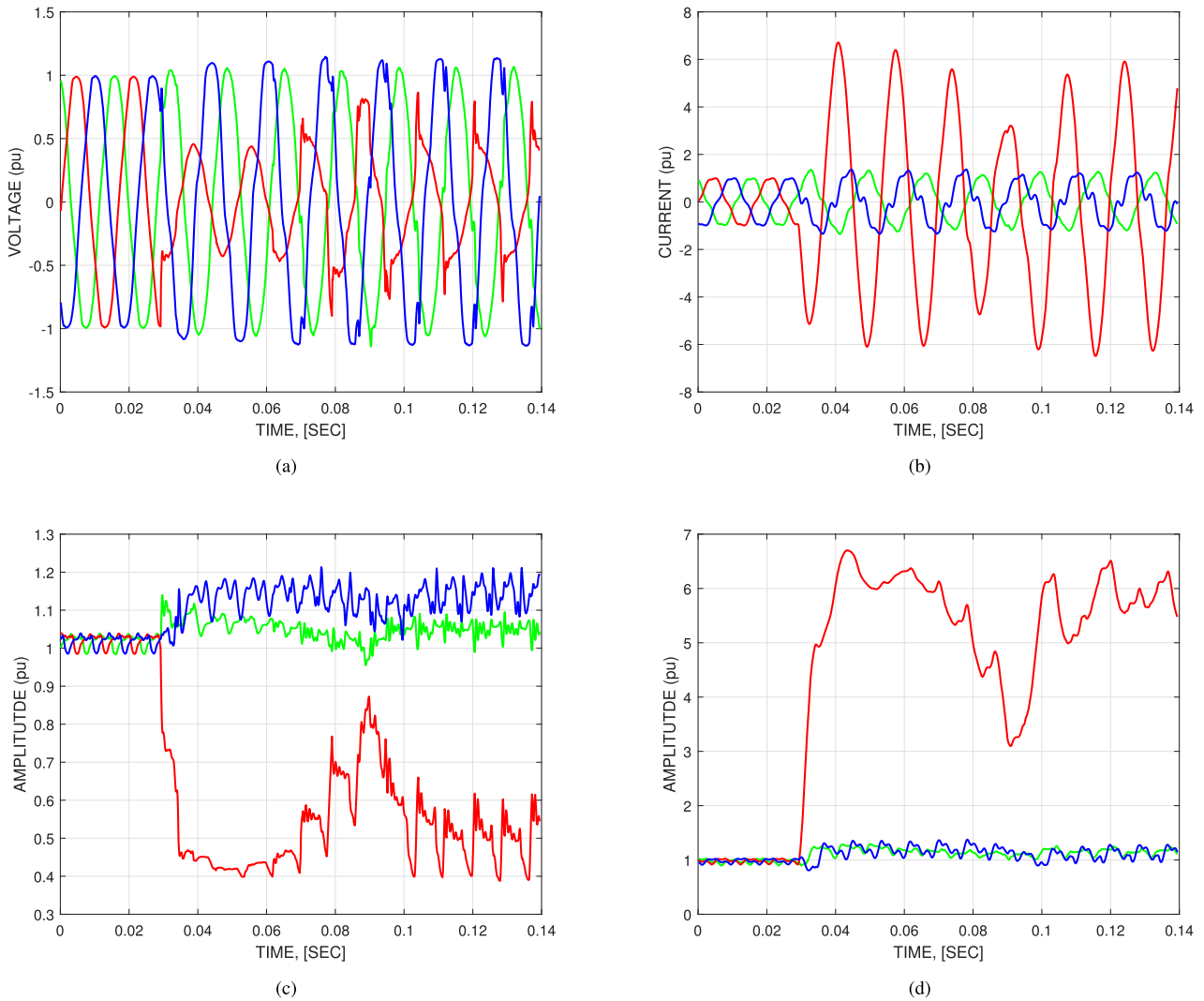


Figure 2. The voltages and the currents which correspond to the sag and swell events in phases ‘a’, ‘b’ and ‘c’ are plotted with green, red and blue lines respectively in (a,b). The events are associated with the fault on the lead cable caused by circuit breaker to trip on substation. The data are obtained from the DOE/EPRI National Database of Power System Events, [26], event ID is 2827. Amplitudes of the first harmonic (with the colours corresponding to the signals in Figures 2(a,b)) which show detectability of the sag and swell events are plotted in Figures 2(c,d).

Notice that recursive Richardson iteration (in comparison with nonrecursive counterpart) is preferable for accuracy critical signal processing applications in noisy environment and in finite digit calculations. For example, the recursive Richardson iteration is preferable for the frequency estimation applications in microgrids (which is both accuracy and time critical application) where the frequency should be estimated with very high accuracy [21,28].

7. Conclusions and future challenges

The family of Richardson algorithms was extended in this paper with (a) new unified framework for recursive convergence accelerators and error models and (b) new nonrecursive algorithms. The paper showed that the Richardson approach represents general optimizable tool which is applicable to many practical problems.

Recursive and nonrecursive Richardson algorithms together with the standard LU decomposition method were considered in practical problem of the detection of the sag and swell signatures in voltage and current signals in three-phase power systems. Real data obtained from the DOE/EPRI National Database of Power System Events were used for a comparison of the algorithms. New nonrecursive Richardson algorithms which save close to half of the computational time compared to the LU decomposition method were recommended for power quality monitoring applications.

Additional significant distortions of voltage and current signals are expected in the future electricity networks due to higher penetration level of renewable energy sources, power electronics, non-linear controllable loads and many others. This will result in challenging detection/estimation problems associated with numerical and computational methods for large-scale

systems, where algorithms developed in this paper can be applied.

Disclosure statement

No potential conflict of interest was reported by the author.

References

- [1] Björck Å. Numerical methods for least squares problems. 1st ed. Philadelphia (PA): SIAM; 1996 April 1.
- [2] Ljung L. System identification: theory for the user. Upper Saddle River (NJ): Prentice-Hall; 1999.
- [3] Stotsky A. Accuracy improvement in least-squares estimation with harmonic regressor: new preconditioning and correction methods, 54-th CDC, Dec. 15–18, Osaka, Japan, 2015. p. 4035–4040.
- [4] Stotsky A. Unified frameworks for high order Newton-Schulz and Richardson iterations: a computationally efficient toolkit for convergence rate improvement. *J Appl Math Comput.* 2019;60(1–2):605–623.
- [5] Dubois D, Greenbaum A, Rodrigue G. Approximating the inverse of a matrix for use in iterative algorithms on vector processors. *Computing.* 1979;22:257–268.
- [6] Stotsky A. Combined high-order algorithms in robust least-Squares estimation with harmonic regressor and strictly diagonally dominant information matrix. *Proc IMechE Part I J Syst Control Eng.* 2015;229(2):184–190.
- [7] Srivastava S, Stanimirovic P, Katsikis V, et al. A family of iterative methods with accelerated convergence for restricted linear system of equations. *Mediterr J Math.* 2017;14–222:1–26.
- [8] Stotsky A. Efficient iterative solvers in the least squares method, IFAC PapersOnLine 53-2, 2020, p. 883–888, see also arXiv:2008.11480v1 [math.OC], 26 Aug 2020.
- [9] Schulz G. Iterative berechnung der reziproken matrix. *Z Angewandte Math Mech.* 1933;13:57–59.
- [10] Hotelling H. Some new methods in matrix calculation. *Ann Math Stat.* 1943;14:1–34.
- [11] Demidovich B, Maron I. Basics of numerical mathematics. Moscow: Fizmatgiz; 1963 (in Russian).
- [12] Ben-Israel A. A note on an iterative method for generalized inversion of matrices. *Math Comput.* 1966;20:439–440.
- [13] Isaacson E, Keller H. Analysis of numerical methods. New York: John Wiley & Sons; 1966.
- [14] Petryshyn W. On generalized inverses and on the uniform convergence of $(I - \beta K)^n$ with application to iterative methods. *J Math Anal Appl.* 1967;18:417–439.
- [15] Stickel E. On a class of high order methods for inverting matrices. *Z Angew Math Mech.* 1987;67:331–386.
- [16] Durand E. Solutions Numeriques des Equations Algebriques. Vol. 2. Paris: Masson; 1972.
- [17] Brezinski C. Variations on Richardson's method and acceleration. In: Numerical Analysis, A Numerical Analysis Conference in Honour of Jean Meinguet. Belgium: Bull Soc Math.; 1996. p. 33–44.
- [18] O'Leary D. Yet another polynomial preconditioner for the conjugate gradient algorithm. *Linear Algebra Appl.* 1991;154–156:377–388.
- [19] Srivastava S, Gupta D. A higher order iterative method for $A_{T,S}^{(2)}$. *J Appl Math Comput.* 2014;46(1/2):147–168.
- [20] Soleymani F, Stanimirovic P, Haghani F. On hyperpower family of iterations for computing outer inverses possessing high efficiencies. *Linear Algebra Appl.* 2015;484:477–495.
- [21] Stotsky A. Towards accurate estimation of fast varying frequency in future electricity networks: the transition from model-free methods to model-based approach. *J Syst Control Eng.* 2016;230(10):1164–1175.
- [22] Zakarya Oubrahim. On electric grid power quality monitoring using parametric signal processing techniques, Universite de Bretagne occidentale – Brest. 2017. English. NNNT: 2017BRES0102.
- [23] Jin T, Chen Y, Guo J, et al. An effective compensation control strategy for power quality enhancement of unified power quality conditioner. *Energy Rep.* 2020;6:2167–2179.
- [24] Liu Y, Jin T, Mohamed MA, et al. A novel three-step classification approach based on time-dependent spectral features for complex power quality disturbances. *IEEE Trans Instrum Meas.* 2021;70:3000814. DOI:10.1109/TIM.2021.3050187
- [25] Naidoo R, Pillay P. A new method of voltage sag and swell detection. *IEEE Trans Power Delivery.* 2007 April;22(2):1056–1063.
- [26] DOE/EPRI national database repository of power system events. Available from: <https://pqmon.epri.com>.
- [27] Artidiello S, Cordero A, Torregrosa J, et al. Generalized inverses estimations by means of iterative methods with memory. *Mathematics.* 2020;8(2):1–13.
- [28] Stotsky A. Frequency regulation in smart microgrids based on load estimation. *J Control Autom Electric Syst Springer.* 2018 Aug;29(4):525–533.