

A NOTE ON WINDOWING FOR THE WAVEFORM RELAXATION

*Hong Zhang*¹

Institute for Computer Applications in Science and Engineering
NASA Langley Research Center
Hampton, VA 23681-0001

ABSTRACT

The technique of windowing has been often used in the implementation of the waveform relaxations for solving ODEs or time dependent PDEs. Its efficiency depends upon problem stiffness and operator splitting. Using model problems, the estimates for window length and convergence rate are derived. The effectiveness of windowing is then investigated for non-stiff and stiff cases respectively. It concludes that for the former, windowing is highly recommended when a large discrepancy exists between the convergence rate on a time interval and the ones on its subintervals. For the latter, windowing does not provide any computational advantage if machine features are disregarded. The discussion is supported by experimental results.

¹Research was supported by the National Aeronautics and Space Administration under NASA Contract No. NAS1-19480 while the author was in residence at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA 23681. Research was also supported in part by the National Science Foundation under contract CCR-9103296.

1 Introduction

The waveform relaxation (WR) method was originally proposed for solving ordinary differential equations arising from very large scale integration (VLSI) circuit simulation [5] [9]. Unlike conventional timestepping methods, it iteratively partitions a big system into mutually decoupled subsystems, and then solves each subsystem independently. Different discretizations and time steps are allowed for integrating subsystems. Based on its nature, the method has been proposed as a multirate method for sequential computing or a parallel method on advanced computers [2].

Under reasonable assumptions for an ODE, the WR iteration has been shown to converge superlinearly on finite intervals [1] [5] [8]. The uniform convergence on an interval of $[0, T]$ is reached in the exponential norm

$$\|u\|_{\xi, T} := \max_{t \in [0, T]} |e^{-\xi t} u(t)|, \quad \xi > 0,$$

which implies that, for many problems, the WR iteration will converge much faster on short intervals than on longer ones. In order to accelerate the convergence, the technique of windowing is recommended, in which the interval of integration is split into a series of subintervals, called *windows*, with iteration taking place on each window successively.

The length of window is of practical important and strongly depends upon problem and machine involved. The general guidance for its selection and the way for evaluating its effectiveness are relatively unknown even though windowing has been a common practice in using the WR method.

The estimates for time windows of the WR iteration were studied by Leimkuhler and Ruehli for RC circuits arising as simplified models of a VLSI interconnect [4]. Finer estimates were developed by Leimkuhler for a model linear second-order system [3]. Using the speed of splitting and weighted spectral radius of iteration operator, Leimkuhler estimated the abscissa of ω -convergence, which then provided a priori estimate for the length of a window wherein convergence was approximately geometric with the given rate ω . His approach puts emphasize on the qualitative comparisons between splittings.

In this paper, we focus on the time dependency of the approximation error and intimate relation between the WR iteration (or dynamic iteration called in [6]) for time dependent problems and the static iteration for corresponding steady-state (or static) problems. For certain model problems, it is

possible to separate the factor that represents early sweeps from the one that dominates asymptotic behavior. Simple convergence estimates are therefore obtained. The estimates and the results observed in the experiments are compared and shown to have good agreements. Based on these estimates, the effectiveness of windowing is discussed for non-stiff and stiff problems respectively. General guidance for the use of windowing is concluded in the end.

2 Waveform relaxation

Using a first-order linear system

$$\frac{du}{dt} + Lu = f, \quad t > 0, \quad u(0) = u_0, \quad (1)$$

with a given splitting $L = M - N$, the WR iteration can be illustrated by

$$\frac{du^{(v)}}{dt} + Mu^{(v)} = Nu^{(v-1)} + f, \quad t > 0, \quad u^{(v)}(0) = u_0.$$

It is an iterative process on a space of differential functions. The functions $u^{(v)}$, so called “waveforms,” are then discretized for numerical integration. The continuous approximation error $e^{(v)} := u^{(v)} - u$ satisfies

$$e^{(v)}(t) = \mathcal{S}e^{(v-1)}(t), \quad t \geq 0, \quad (2)$$

where \mathcal{S} is a linear operator on $L^p(R^+, C^n)$ ($1 \leq p \leq \infty$) depending upon M and N , and is called the *iteration operator*.

Several convergent splittings for an ODE system were proposed and discussed in [3] and [6]. We shall restrict ourselves to splittings that resemble Jacobi splitting and Gauss-Seidel splitting on linear systems with time independent coefficients, as described by Eq.(1). For simplicity, the space considered in this work is $C^\infty([0, T], C^n)$, the space of continuous C^n -valued functions in $[0, T]$, with $\|\cdot\|$ denotes l^∞ norm for space variables:

$$\|u(t)\| := \max_{1 \leq i \leq n} |u_i(t)|,$$

and $\|\|\cdot\|\|_T$ stands for

$$\|\|u\|\|_T := \max_{t \in [0, T]} \|u(t)\|.$$

The notation $\|\cdot\|$ will also be used as l^∞ induced matrix and operator norm. The issue of time discretization is beyond our consideration, for which the reader may refer to [7].

3 Convergence estimates

For certain type of problems or operator splittings, the factor representing early phase of iterations and the factor dominating asymptotic behavior in the approximation error can be separated. Laplace transform is a convenient tool for doing this. Applying Laplace transform to Eq.(2), the error is expressed as

$$\hat{e}^{(v)}(z) = S(z)\hat{e}^{(v-1)}(z) = S^v(z)\hat{e}^{(0)}(z), \quad \operatorname{Re}z \geq 0,$$

where

$$S(z) = (zI + M)^{-1}N$$

is the Laplace transform of the convolution kernel of \mathcal{S} . Note, $S(0)$ is the iteration operator for the steady-state problem $Lu = f$ corresponding to Eq.(1).

Theorem 1. Let L be split as $L = M - N$ with $M = dI$, $d > 0$. After v WR iterations, the error is bounded by

$$\|e^{(v)}(t)\| \leq g_v(dt) \cdot \|S^v(0)\| \cdot \|e^{(0)}\|_T, \quad t \in [0, T],$$

where

$$g_v(t) = \frac{\Gamma_t(v)}{\Gamma(v)} = 1 - e^{-t} \left(\sum_{i=0}^{v-1} \frac{t^i}{i!} \right) = e^{-t} \sum_{i=v}^{\infty} \frac{t^i}{i!}, \quad (3)$$

and $\Gamma_t(v)$ is the incomplete Γ -function.

Proof. For $M = dI$, $d > 0$,

$$S^v(z) = ((z + d)^{-1}N)^v = (z/d + 1)^{-v} S^v(0).$$

Let

$$\hat{f}_v(z) := (z + 1)^{-v}.$$

The inverse Laplace transform of \hat{f}_v is

$$f_v(t) = \frac{1}{(v-1)!} e^{-t} t^{v-1}.$$

Hence the error in the time domain satisfies

$$e^{(v)}(t) = S^v(0) \int_0^{dt} f_v(\tau) e^{(0)}(t - \tau/d) d\tau.$$

Define

$$g_v(t) := \int_0^t f_v(\tau) d\tau = \frac{\Gamma_t(v)}{\Gamma(v)}, \quad t \geq 0.$$

For $t \in [0, T]$,

$$\begin{aligned} \|e^{(v)}(t)\| &\leq \|S^v(0)\| \cdot \max_{1 \leq i \leq n} \left(\int_0^{dt} f_v(\tau) d\tau \cdot \max_{t \in [0, T]} |e_i^{(0)}(t)| \right) \\ &= g_v(dt) \cdot \|S^v(0)\| \cdot \|e^{(0)}\|_T. \end{aligned}$$

Equation (3) can be easily verified by induction. \square

It is interesting to examine the bounds given by Theorem 1. Note that $(\|S^v(0)\|)^{1/v}$ is time independent and approximates the asymptotic convergence rate either for obtaining the steady-state solution or solutions over long time intervals. Function g_v represents the time dependency of the error and dominates the convergence behavior at early phase of iteration or on short intervals. It is a monotone increasing function, bounded by $0 \leq g_v(t) < 1$ with

$$g_v(0) = 0, \quad \lim_{t \rightarrow \infty} g_v(t) = 1^-.$$

These observations on the error bound agrees with computational experiences that the WR iteration converges faster on short intervals than on longer ones, and the convergence rates on any time intervals, including infinite interval, are at least as good as the one for the static iteration.

When operator L has constant diagonal, Theorem 1 actually gives an error bound for Jacobi WR iteration. The next two theorems will give results for Gauss-Seidel WR iteration on model problems.

Theorem 2. Let $L = M - N$. Assume that M and N are simultaneously diagonalizable by matrix X , and all eigenvalues of M are positive. Then

$$\|e^{(v)}(t)\| \leq g_v(dt) \cdot \|S^v(0)\| \cdot \text{cond}(X) \cdot \|e^{(0)}\|_T, \quad t \in [0, T],$$

where d is the largest eigenvalue of M and $\text{cond}(X) = \|X^{-1}\| \cdot \|X\|$.

Proof. From the assumptions, there are diagonal matrices Λ_M and Λ_N , such that

$$M = X^{-1}\Lambda_M X, \quad N = X^{-1}\Lambda_N X,$$

and

$$\Lambda_M = \{\lambda_i(M)\}, \quad \lambda_i(M) > 0 \text{ for all } i.$$

In the Laplace domain,

$$S^v(z) = X^{-1}((zI + \Lambda_M)^{-1}\Lambda_N)^v X = (X^{-1}\Lambda_M^{-v}\Lambda_N^v X)(X^{-1}(z\Lambda_M^{-1} + I)^{-v} X),$$

leading to

$$\hat{e}^{(v)}(z) = S^v(0)X^{-1} \begin{bmatrix} \hat{f}_v(z/\lambda_1(M)) & & & \\ & \ddots & & \\ & & \hat{f}_v(z/\lambda_i(M)) & \\ & & & \ddots \end{bmatrix} X \hat{e}^{(0)}(z).$$

Applying inverse Laplace transform, one obtains

$$\begin{aligned} \|e^{(v)}(t)\| &\leq \|S^v(0)\| \cdot \|X^{-1}\| \cdot \max_{1 \leq i \leq n} \left| \int_0^{\lambda_i(M)t} f_v(\tau) \sum_{j=1}^n x_{ij} e_j^{(0)}(t - \tau/\lambda_i(M)) d\tau \right| \\ &\leq \|S^v(0)\| \cdot \|X^{-1}\| \cdot \max_{1 \leq i \leq n} (g_v(\lambda_i(M)t) \sum_{j=1}^n |x_{ij}|) \cdot \|e^{(0)}\|_T \\ &\leq g_v(dt) \cdot \|S^v(0)\| \cdot \text{cond}(X) \cdot \|e^{(0)}\|_T. \quad \square \end{aligned}$$

The lexicographic (or forward point) Gauss-Seidel WR iteration on heat equation with periodic boundary condition is an example that this theorem can be applied (see [10]). When operator L has the form

$$L = d \begin{bmatrix} I & -B \\ -R & I \end{bmatrix}, \quad d > 0, \quad (4)$$

a common structure when using a red-black ordering on certain model problems in ODEs or in time dependent PDEs, the error in Gauss-Seidel WR has a similar upper bound.

Theorem 3. For operator L represented by Eq.(4), the error after v Gauss-Seidel WR iterations, expressed as $e^{(v)}(t)^T = [e_R^{(v)}(t)^T, e_B^{(v)}(t)^T]$, is bounded by

$$\|e^{(v)}(t)\| \leq g_{2v-1}(dt) \cdot \|S^v(0)\| \cdot \|e_B^{(0)}\|_T, \quad t \in [0, T]. \quad (5)$$

Proof. For Gauss-Seidel splitting,

$$M = d \begin{bmatrix} I & 0 \\ -R & I \end{bmatrix}, \quad \text{and } N = d \begin{bmatrix} 0 & B \\ 0 & 0 \end{bmatrix}.$$

Laplace transform of the convolution kernel of \mathcal{S} can be expressed as

$$S(z) = (zI + M)^{-1}N = \begin{bmatrix} 0 & \frac{1}{z/d+1}B \\ 0 & \frac{1}{(z/d+1)^2}P \end{bmatrix}, \quad P = RB, \quad \text{Re}z \geq 0,$$

which yields

$$S^v(z) = \begin{bmatrix} 0 & \hat{f}_{2v-1}(z/d)BP^{v-1} \\ 0 & \hat{f}_{2v}(z/d)P^v \end{bmatrix}$$

and

$$\hat{e}^{(v)}(z) = \begin{bmatrix} \hat{f}_{2v-1}(z/d)BP^{v-1}\hat{e}_B^{(0)}(z) \\ \hat{f}_{2v}(z/d)P^v\hat{e}_B^{(0)}(z) \end{bmatrix}.$$

Back to the time domain,

$$e^{(v)}(t) = \begin{bmatrix} BP^{v-1} \int_0^{dt} f_{2v-1}(\tau)e_B^{(0)}(t - \tau/d)d\tau \\ P^v \int_0^{dt} f_{2v}(\tau)e_B^{(0)}(t - \tau/d)d\tau \end{bmatrix}.$$

Using the properties

$$g_{v+1}(t) < g_v(t), \quad t > 0, \quad v > 0,$$

and

$$\max\{\|BP^{v-1}\|, \|P^v\|\} \leq \|S^v(0)\|,$$

(note, $S^v(0) = \begin{bmatrix} O & BP^{v-1} \\ O & P^v \end{bmatrix}$), the inequality (5) follows immediately. \square

The discussion above indicates that the error at early stage of the WR iteration is controlled by function g_v . Simple convergence estimates are therefore derived naturally from this function. Given a convergence rate ω and an iteration number v , the length of a window of ω -convergence (see [3]) can be estimated by

$$T_\omega := \max\{t : (g_{mtd(v)}(dt))^{1/v} \leq \omega\}, \quad (6)$$

where $mtd(\cdot)$ is an integer function, defined by the operator splitting. For instance, $mtd(v) = v$ and $mtd(v) = 2v - 1$ for the cases discussed in Theorem 1 and Theorem 3 respectively. A useful variant of Eq.(6) is the value

$$\omega_T := (g_{mtd(v)}(dT))^{1/v}, \quad (7)$$

which gives an estimated average convergence rate on windows of length T in first v sweeps.

Example 1. Consider the ODE system

$$\frac{du}{dt} + Lu = 0,$$

where $L = [-1, 2, -1]$, a symmetric tridiagonal matrix with 2 and -1 on main and off diagonal. This system describes the nodal voltage of a linear resistor-capacitor (RC) network [3] [4]. Jacobi WR iteration was performed with randomly generated starting function $u^{(0)}$. The trapezoidal rule was used in the integration with conservative time step $\Delta t = 0.01$ for simulating time-continuous iteration. For a given number of sweeps v , the observed convergence rates ω_{obs} were collected as

$$\left(\max_{1 \leq i \leq n} \max_{t \in [0, T]} \left| \frac{u_i^{(v)}(t) - u_i^{(v-1)}(t)}{u_i^{(v)}(t)} \right| \right)^{1/v}.$$

Figure 1 depicts the graph of T_ω for 5 Jacobi iterations ($v = 5$) together with observed data marked by x's. Table 1 shows more detailed comparison between the observed convergence rate ω_{obs} and estimated rate ω_T on interval

$[0, T]$. The estimates T_ω 's and ω_T 's provided by Eq.(6)-(7) are surprisingly close to the observed data. Although such a good agreement cannot be predicted in general, reasonable match at early phase of iteration should be expected if similar error bounds, as given by these theorems, are conjectured.

Figure 1: Observed and estimated window length

Table 1: Observed and estimated convergence rate on $[0, T]$

v	ω_{obs}/ω_T				
	T=0.2	T=0.4	T=0.6	T=0.8	T=1.0
3	.2271/.1994	.3840/.3620	.4920/.4939	.5682/.6006	.6227/.6864
5	.1536/.1437	.2793/.2691	.3801/.3783	.4622/.4730	.5295/.5550
9	.0000/.0927	.1796/.1781	.2555/.2568	.3242/.3292	.3852/.3956

4 Effectiveness of windowing

The theoretical analysis and practical experiments have revealed that the effectiveness of the WR method depends highly on the stiffness of the ODE solved. Not surprisingly, the effectiveness of using windows in the WR iteration is also closely related to the stiffness of the system. In this section, we investigate their relation, and show how to estimate the efficiency of windowing, which then results the general guidance for the implementation of the WR iteration.

The effectiveness of windowing is discussed in terms of computational cost or operation counts. Following concepts and notations are needed.

Let μ be average operation counts on unit windows per sweep. For a given error tolerance $\epsilon > 0$, $v_T(\epsilon)$ denotes the average number of iterations needed for the convergence on windows of length T . Using Theorem 1, $v_T(\epsilon)$ can be estimated as

$$\tilde{v}_T(\epsilon) := \min\{v : g_{mtd(v)}(dT) < \epsilon\}.$$

The error tolerance ϵ will be dropped whenever the context is clear. The total computations on a window of length T for the convergence is then approximately equal to

$$C := v_T T \mu.$$

Note that $T\mu$ is the average computations on a window of length T per sweep. Let this window be split into two subwindows of length T_1 and T_2 , $T_1 + T_2 = T$, with the WR iteration taking place on each of them until the tolerance level is reached one after the other. The total cost then satisfies

$$v_{T_1} T_1 \mu + v_{T_2} T_2 \mu \leq v_{T'} T \mu, \quad T' = \max\{T_1, T_2\}.$$

Define $C_{win} := v_{T'}T\mu$. C and C_{win} will be referred as the average computations on an interval of length T without and with windowing respectively. Since $v_{T'} \leq v_T$, we have

$$C_{win} \leq C. \quad (8)$$

This indicates that windowing does not introduce extra computation.

The number of iterations for the convergence on an interval of length T is closely related to the stiffness of the ODE involved. This can be seen from the behavior of the function g_v (see Eq.(3)). For example, an approximation for the case stated in Theorem 1 is

$$g_v(dT) \approx \frac{(dT)^{\tilde{v}_T}}{\tilde{v}_T!} \approx \epsilon,$$

leading to

$$\epsilon^{1/\tilde{v}_T} (\tilde{v}_T/2)^{1/2} < dT \approx (\epsilon \tilde{v}_T!)^{1/\tilde{v}_T} < \epsilon^{1/\tilde{v}_T} \tilde{v}_T.$$

That is, the estimated number of iterations \tilde{v}_T increases proportional to the parameter d , an indicator of the stiffness of the ODE solved.

4.1 Non-stiff case

Given an interval of length T . Following above arguments, the number of the WR iterations for convergence is not too large for non-stiff systems. If the interval is split into k windows of equal length $T' = T/k$, the gain or the percentage of the savings of using windows can be measured by

$$\begin{aligned} S_{win} : &= \frac{C - C_{win}}{C} = \frac{v_T T \mu - k v_{T'} T' \mu}{v_T T \mu} \\ &= \frac{v_T - v_{T'}}{v_T}. \end{aligned}$$

If $v_T \approx v_{T'}$, $S_{win} \approx 0$, not much computation can be saved by windowing. When $v_T > v_{T'}$, using windows, $v_T - v_{T'}$ sweeps of WR on this interval of length T are likely to be reduced.

Table 2 lists the experimental results on Example 1 discussed in Section 3, as well as corresponding estimated values. The entries are of the form *observed/estimated*. The numbers listed inside parenthesis in column 2 are average number of iterations collected on subintervals of $[0, 2]$ with length T . Windowing was used on intervals $[0, T]$, $T = 0.5, 1.0$, and 2.0 with the

window length 0.25. The iteration was terminated when the relative error was in the order of $\epsilon = 1.e - 7$. Again, the estimates were in good agreement with the observed ones. Windowing reduced the computations by 25-60%. It clearly suggests that windowing is quite efficient in reducing computations when a large discrepancy exists between the convergence rate on an interval and the ones on its subintervals.

Table 2: Effectiveness of windowing (observed/estimated)

Interval [0, T]	No. of Iterations $v_T(\text{ave. } v_T)/\tilde{v}_T$	Computations		S_{win} %
		Without Windowing	With Windowing	
[0, 0.25]	8 (7) /8	$2\mu / 2\mu$	$2\mu / 2\mu$	
[0, 0.50]	10 (9) /10	$5\mu / 5\mu$	$3.75\mu / 4\mu$	25 / 20
[0, 1.00]	13 (12) /13	$13\mu / 13\mu$	$7.25\mu / 8\mu$	44 / 38
[0, 2.00]	17 (17) /18	$34\mu / 36\mu$	$13.5\mu / 16\mu$	60 / 56

Remark. Eq.(8) and the observed data in Table 2 seem to suggest choosing minimum window length, which ironically is equivalent to the step size used in the time integration of subsystems. In this situation, the WR method is nothing but a time-stepping method for solving ODEs. However, recall that the method is proposed as a multirate method in the context of serial computation or a parallel method on advanced computers. For the former, it is developed for problems in which the coupling of subsystems is relatively loose and many subsystems allow large integration steps. The window length is therefore recommended as the largest step size used in the time integration of subsystems. For the later, machine characters such as vector length, communication overhead, play important roles. The study in this paper is restricted to the mathematical concerns only.

4.2 Stiff case

In this situation, the WR iteration would take large number of iterations to converge in an interval. The convergence rate very likely has entered the asymptotic behavior. Estimating it in terms of function g_v alone is no longer

adequate. From the error bounds given by Section 3, the rate of convergence at time t would be dominated by

$$(g_{mtd(v)}(dt)||S^v(0)||)^{1/v} \approx \rho(S(0)),$$

the spectral radius of $S(0)$ or the convergence rate of the related static iteration. Since $\rho(S(0))$ is time independent, the convergence rates on any intervals are almost identical, so are the numbers of iterations needed for the convergence on those intervals.

Example 2. Consider the heat equation on the unit square $\Omega = (0, 1) \times (0, 1)$ with Dirichlet boundary conditions

$$\begin{aligned} u_t - \Delta u &= 0, & (t, x) \in (0, T] \times \Omega \\ u &= 0, & (t, x) \in [0, T] \times \partial\Omega, \quad u(0, x) = u_0(x), \quad x \in \Omega. \end{aligned}$$

The equation was first discretized in space, resulting the semi-discrete problem

$$\frac{dU}{dt} + L_h U = 0, \tag{9}$$

where L_h is the five point difference approximation operator to the Laplacian

$$L_h := \frac{1}{h^2} \begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix},$$

and h is the mesh size of space discretization. The red-black Gauss-Seidel WR was then implemented on the system (9) over time interval $[0, T]$. The iteration was terminated when the difference between the v th and $v - 1$ th approximation

$$|||U^{(v)} - U^{(v-1)}|||_T$$

reached the level of truncation error $O(h^2)$, a safe stopping criterion proven by Nevanlinna [8]. Table 3 shows that the numbers of iterations needed on different time intervals are almost the same, confirming the above arguments. As is discussed, using small windows for this problem virtually has no mathematical advantage.

Table 3: Number of iterations

h	[0, 0.125]	[0, 0.25]	[0, 0.5]	[0, 1.0]
1/8	16	17	17	17
1/16	53	60	60	60
1/32	208	239	239	239

5 Conclusions

In this paper, the convergence estimates such as window length and convergence rate are developed using the qualitative comparison between the WR iteration and the corresponding static iteration. The effectiveness of windowing technique for the WR method is discussed. The results proven in this work and observed in the experiments suggest that, for non-stiff ODEs, substantial computations can be saved by windowing when a large discrepancy exists between the convergence rate on an interval and the ones on its subinterval; while for stiff problems, which are typically arisen from time dependent PDEs, windowing has no mathematical advantage. Thus for stiff systems, the selection of the window length should be mainly determined by the machine features, such as memory capacity, vector length, cache size, communication cost, etc.

Although only a few model problems are considered in this work, similar approach could be taken for some generalized problems. The guidance concluded above certainly provide helpful information in the implementation of the WR method for wide class of applications.

References

- [1] A. BELLEN AND M. ZENNARO, *The use of Runge-Kutta formulae in waveform relaxation methods*, Appl. Numer. Math., 11 (1993), pp. 95–114.
- [2] C. GEAR, *Waveform methods for space and time parallelism*, J. of Computational and Applied Mathematics, 38 (1991), pp. 137–147.

- [3] B. LEIMKUHNER, *Estimating waveform relaxation convergence*, SIAM J. Sci. Comput., 14 (1993), pp. 872–889.
- [4] B. LEIMKUHNER AND A. RUEHLI, *Rapid convergence of waveform relaxation*, Appl. Numer. Math., 11 (1993), pp. 211–224.
- [5] E. LELARSMEE, A. RUEHLI, AND A. SANGIOVANNI-VINCENTELLI, *The waveform relaxation method for time-domain analysis of large scale integrated circuits*, IEEE Trans. Computer-Aided Design, 1 (1982), pp. 131–145.
- [6] U. MIEKKALA AND O. NEVANLINNA, *Convergence of dynamic iteration methods for initial value problems*, SIAM J. Sci. Stat. Comput., 8 (1987), pp. 459–482.
- [7] ———, *Sets of convergence and stability regions*, BIT, 27 (1987), pp. 554–584.
- [8] O. NEVANLINNA, *Remarks on Picard Lindelöf iteration*, BIT, 29 (1989), pp. 328–346.
- [9] A. NEWTON AND A. SANGIOVANNI-VINCENTELLI, *Relaxation-based electrical simulation*, IEEE Trans. Computer-Aided Design, 3 (1984), pp. 308–331.
- [10] S. TAÁSAN AND H. ZHANG, *On the waveform multigrid method*, Tech. Report 93-21, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, (1993).