The Holomorphic Embedding Load Flow Method

Antonio Trias

Abstract—**The Holomorphic Embedding Load Flow is a novel general-purpose method for solving the steady state equations of power systems. Based on the techniques of Complex Analysis, it has been granted two US Patents. Experience has proven it is performant and competitive with respect established iterative methods, but its main practical features are that it is non-iterative and deterministic, yielding the correct solution when it exists and, conversely, unequivocally signaling voltage collapse when it does not. This paper reviews the embedded load flow method and highlights the technological breakthroughs that it enables: reliable real-time applications based on unsupervised exploratory load flows, such as Contingency Analysis, OPF, Limit-Violations solvers, and Restoration plan builders. We also report on the experience with the method in the implementation of several real-time EMS products now operating at large utilities.**

Index Terms—**Load flow analysis, power system modeling, power system simulation, power engineering computing, energy management, decision support systems, power system restoration.**

I. Introduction

MOST known load flow methods, particularly those used in software for large real-world networks, are based on general-purpose numerical iterative techniques (for a recent review, see [1]). The earliest methods were based on Gauss-Seidel (GS) iteration [2], which has slow convergence rates but very small memory requirements. GS may still be used when the other methods fail to converge starting from the flat profile, but methods based on Newton-Raphson (NR) [3] are judged to be better in general, because of their quadratic convergence properties. However, even after applying sparse linear algebra techniques for efficiency [4], the full NR method is computationally too expensive due to the reevaluation of the Jacobian matrix and the corresponding factorizations that need to be done at each iteration. Several techniques exploit the weak coupling between active power and voltage magnitude on the one hand, and reactive power and phase angles on the other, in order to derive more efficient schemes [5], [6] where the Jacobian matrix only needs to be factorized once. Of all the various decoupled methods based on NR, the so-called Fast Decoupled Load Flow (FDLF) formulation of Stott and Alsac [7] has become the most successful and it is almost a de-facto standard in the industry, either in its original form or in one of its variants [8].

To a greater or lesser degree, all these iterative methods suffer from the same convergence pitfalls: on the one hand, there is no guarantee that the iteration will always converge, as this depends on the choice of the initial point; on the other, since the system has multiple solutions, it is not always possible to control to which solution it will converge. As it is well-known the load flow equations have multiple solutions, and only one of them corresponds to the real operative state of the electrical system. Unless one provides a starting point sufficiently close to the correct solution, iterative schemes may not just fail to converge, but converge to a spurious solution.

Although these problems are well known, there is a certain lack of awareness among practitioners. The reasons are clear: under normal operating conditions, convergence problems do not appear often, and even the flat profile provides a good initial seed. Some heuristics have been devised to come up with better starting seeds [9], [10], which helps to minimize non-convergence cases, and efforts have been made to understand and characterize the regions of convergence [11]. In any case, non-convergence is mostly a problem in real-time, where there is no time to manually tune the seed until convergence is reached, or, more critically, to review the solution in detail to check for possible spurious convergence at a few buses. In the author's opinion, it is not sufficiently recognized how these reliability problems have hindered the development of real-time applications for the operators of power transmission and distribution networks. Intelligent real-time applications have a tremendous potential in power systems because, in contrast to other infrastructures under human control, there is an underlying physical model that is for the most part completely descriptive and accurate. However, real-time usage requires that the physical model be solved in a fully 100% deterministic and reliable way. The Holomorphic Embedding load flow method is posed to fill in this important gap.

The rest of the paper is organized as follows. Section II describes the problems of convergence of iterative methods and discusses the contexts in which such problems become relevant. Section III exposes the Holomorphic Embedding Load Flow Method (HELM), from the concepts of complex embedding to the practical mechanics of the method. Section IV exemplifies the method step by step, on the two-bus model. Finally, Section V briefly reports on the real-world experience in the development of intelligent real-time applications that benefit from the full reliability of the method.

II. The problem of convergence in iterative methods

Numerical iterative methods are widely used in numerical analysis for the solution of nonlinear equations. But contrary to linear systems, where convergence is ensured and unique, nonlinear systems exhibit a wide array of potential problems. Convergence is only ensured when the initial seed is an estimate "sufficiently close" to the actual solution (so that the conditions for the Banach fixed-point theorem are satisfied), but there are no practical criteria for quantifying how close this needs to be in general. Then systems having multiple

A. Trias is with Aplicaciones en Informatica Avanzada SA, Sant Cugat del Vallès, Spain. e-mail: triast@aia.es

Manuscript received November 30, 2011. Reviewed February 8, 2012.

978-1-4673-2729-9/12/$31.00 ©2012 IEEE

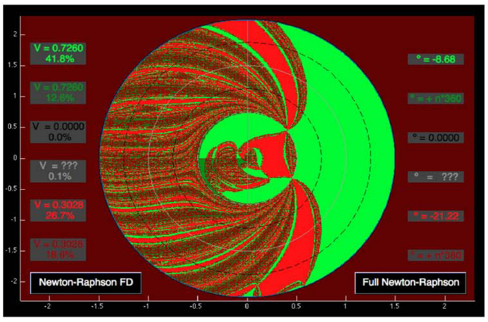

Fig. 1. Fractal basins of attraction for the two-bus load flow problem, under the FDLF method. Initial seeds that lead to the correct solution are shown in green; to the spurious solution in red; and non-convergence in black.

solutions pose additional problems, since the iterations may converge to a non-desired solution. When different solutions get too close to each other, there are no general methods to deterministically select the desired one, unless one can exploit particular symmetries in the system.

Moreover, for systems in complex variables (as it is the case of the load flow problem) the boundaries between the basins of attraction under the iterative scheme are typically fractal. As two or more solutions get closer, their respective basins become more and more intertwined, so that, eventually, the neighborhood of any given solution becomes peppered by points attracted to different ones. In the case of the Newton-Raphson method, this phenomenon has been studied extensively and has been given the name of *Newton fractals* [12]. It is a straightforward exercise to verify that Newton fractals are indeed present in the load flow problem, using either the full NR or the FDLF method, as several authors have shown [13], [14]. Figure 1 shows an example using the two-bus model, where the results can be contrasted against the two exact solutions. The figure, which represents the complex plane for the voltage variable V, is obtained simply by coloring each point according to where the iteration leads to. Typically, one would assign different colors for at least these three possible outcomes: the correct load flow solution, the spurious low-voltage solution, and no-convergence. One could also use colors for convergence to non-solutions (more rare, but it also happens), and for other numerical artifacts such as the convergence of phase angles on high multiples of 2π. The author provides more examples on the web site [15].

For the load flow problem the consequence is that, for arbitrary electrical scenarios, it is hopelessly difficult to devise any *a priori* mechanism to select a good seed for the iterative method. Some efforts have been devoted to improve the seed [9], [10], but they cannot guarantee success. The problem becomes progressively more evident as the network load increases: when one bus approaches its maximum loadability point (voltage collapse), a spurious solution gets closer to the real one, in what may be understood as a saddle-node bifurcation scenario [16], [17]. Under these stressed conditions it becomes increasingly difficult for iterative methods to ensure convergence to the correct solution, or any convergence at all.

There is a class of methods, commonly referred to as the Continuation Load Flow [18], which try to circumvent this problem. These use the numerical techniques of homotopic continuation, or "path-following" [19], which allow to faithfully track a known solution around the bifurcation point. The problem is that, in real-time EMS applications, one commonly needs to calculate load flow solutions for hypothetical *what-if* scenarios that are not smoothly connected to any previously known one. This happens for instance in many contingency studies, and most notably in any application that explores the state space by simulating SCADA actions. Therefore using a continuation load flow for real-time applications is fundamentally unfeasible, let alone slow.

III. The Holomorphic Embedded Load Flow Method

As evidenced in the previous section there is a need for a direct, fully reliable load flow method. One theoretical possibility is solving the equations exactly in closed form, using polynomial elimination techniques (resultants and Groebner Basis) with the help of computer algebra packages [20], [21]. This has very strong limitations, since the memory and computational costs make it impossible to get past 5 or 6 buses. A more interesting approach from the practical point of view is the so-called Series Load Flow [22], [23], based on an earlier idea by Sauer [24]. This method uses the Taylor expansion of the voltage variables as functions of all the specified parameters of the problem (power injections, voltage magnitudes), calculated on a point in which the solution is known. Summation of the Taylor series then allows to extend the solution to other scenarios, in a process that is limited by the radius of convergence of the series.

Although it was developed completely independently, the Holomorphic Embedding method presented in this paper is somewhat related to these ideas, but with one key difference: the Series Load Flow uses real variables and therefore cannot guarantee convergence of the Taylor series in general, for arbitrary ranges. By contrast the Holomorphic Embedding Load Flow method described in this paper is based on Complex Analysis. This seemingly minor technicality makes an enormous difference, and not just in terms of gaining new theoretical insights. It is only by working in the complex field and using the wonderful properties of holomorphic functions, that the method achieves its desired properties. In short, the Holomorphic Embedding Load Flow provides a procedure to compute, with mathematically proven guarantees of success, the right solution to the desired accuracy (within the constraints of the computer arithmetic accuracy); and otherwise signals unambiguously if the system has no solution (voltage collapse).

A. *Holomorphic Embedding*

In the following, the main steps of the Holomorphic Embedding method are reviewed. The method is based on the concept of embedding as general problem-solving program: the original problem is first submerged (embedded) in a larger problem, in which the solution is easy to find under some reference conditions. Then one expects to be able to use such reference solution away from those conditions, the goal being to compute the solution to the original problem. In this case the method proposes embedding the original algebraic equations in a holomorphic functional extension of them, which allows us to exploit the nice properties endowed by complex analyticity.

For the sake of clarity, the method will be exposed in the case where all buses are of type PQ. The treatment of PV nodes and other types of controls is touched upon in Section V-A. Consider then the following general form for a load flow problem:

$$\sum_k Y_{ik} V_k = I_i^{\text{load}} + \frac{S_i^*}{V_i^*} \tag{1}$$

where Y_{ik} is the generalized admittance containing branch admittances, bus shunt admittances, and any constant-impedance injections. Symmetry in Y_{ik} is not required, so phase shifting transformers are allowed. The right hand side is left with the constant-injection and constant-power components of a general ZIP load model. The proposed embedding consists in introducing a complex parameter s into (1) so that the voltages become functions of this new complex variable. As it will be shown, it is essential to the method to work in the complex field. The embedding can be done in various ways, but the method explicitly proposes the following form:

$$\sum_k Y_{ik} V_k(s) = s I_i^{\text{load}} + \frac{s S_i^*}{V_i^*(s^*)} \tag{2}$$

This particular embedding satisfies the first requirement of the method: at $s = 0$ all injection terms vanish and the system is trivially solvable by linear algebra. It represents the system under no load or generation, just swing-bus sources propagating voltage everywhere (and in the absence of shunts, the solution is exactly $|V_i| = 1, \theta_i = 0$ everywhere). As it will be shown below, this reference point used by the method can be unambiguously defined.

Secondly, it is required for the embedding to be holomorphic, that is, it should define the voltages V_i to be holomorphic functions in the embedding parameter s. This is done in order to benefit from all the power of complex analysis, in particular the process of analytical continuation that will allow to obtain the objective state from the reference one. It should be strongly remarked that the denominator on the right hand side of (2) has the form $V^*(s^*)$, and not $V^*(s)$. It can be shown that, because complex conjugation does not leave the Cauchy-Riemann equations invariant, this is the only choice that allows $V(s)$ to have a chance of being holomorphic. It is then useful to define $\overline{V}(s) \equiv V^*(s^*)$, so that the embedded system becomes:

$$\sum_k Y_{ik} V_k(s) = s I_i^{\text{load}} + \frac{s S_i^*}{\overline{V}_i(s)}$$

However this change of notation does not hide the fact that is cumbersome to proceed with the analysis of the embedded equations by requiring $\overline{V}_i(s) = V_i^*(s^*)$ throughout the treatment. In order to escape this difficulty, the embedding method proposes to study the following system of algebraic equations:

$$\begin{aligned} \sum_k Y_{ik} V_k(s) &= s I_i^{\text{load}} + \frac{s S_i^*}{\overline{V}_i(s)} \\ \sum_k Y_{ik}^* \overline{V}_k(s) &= s I_i^{*\text{load}} + \frac{s S_i}{V_i(s)} \end{aligned} \tag{3}$$

where $V_i(s), \overline{V}_i(s)$ are now independent complex functions representing the two degrees of freedom of the voltages. Note that when

$$\overline{V}_i(s) = V_i^*(s^*) \tag{4}$$

these equations are just complex conjugates of each other, as expected. However, it should be emphasized that the converse is not true: the condition (4) is not implied by the algebraic system (3), and in fact there may exist solutions which do not satisfy the condition and therefore are not physical solutions to the original load flow (1). Therefore the embedding method consists in solving the algebraic system (3) *and* requiring the additional condition (4), which will be referred to as the reflection condition.

From here on, the following terminology is adopted:

- solutions to the algebraic embedded system which do not satisfy (4) will be referred to as *ghost solutions*. They are simply not a solution of the load flow equations.
- solutions that do satisfy the reflection condition will be referred to as *physical solutions*. As it will be seen below, these can be either "normal" (corresponding to correct operating conditions, of which there may be only one), or "anomalous" (corresponding to unstable operating conditions, of which there may be several, in general). Since anomalous solutions originate in the physical state corresponding to low voltage magnitude (low load impedance), these will be referred to as *black solutions*, and the normal solution will be referred to as the *white solution*.

Up to this point it remains to be shown that the embedding in (3) does in fact define $V_i(s)$ and $\overline{V}_i(s)$ as holomorphic functions. It turns out that, since (3) are algebraic, elimination techniques based on the theory of resultants and Gröbner Basis [25] guarantee that all variables V_i can be successively eliminated in terms of the remaining ones, until a polynomial equation in V_1 is obtained:

$$\mathfrak{P}(V_1) = \sum_{n=0}^{N} p_n(s) V_1^n = 0 \tag{5}$$

and the $\overline{V}_1, V_2, \overline{V}_2, V_3$, etc., are expressed explicitly as polynomials in all the previous ones in a triangular manner, which allows the obtention of all other V_i and $\overline{V}_i$ from each solution V_1, by simple progressive back-substitution. The degree N of this polynomial is in general rather large (of order exponential in the number of variables V_i in the original system), but always finite, and the coefficients $p_n(s)$ are polynomial in s. This is precisely the definition of an algebraic curve. Therefore all V_i and $\overline{V}_i$ are proved to be holomorphic functions everywhere except on a finite number of points, known as the exceptional set of the algebraic curve. These exceptional points are those values of s on which the polynomial equation for V_1 exhibits a null derivative $\partial\mathfrak{P}/\partial V_1$, and will play an important role on the discussion about analytic continuation further below.

B. Power series

Obtaining and solving (5) for large networks is unfeasible in practice. Instead, one should work with the power series of the holomorphic functions, as is commonly done when calculating algebraic curves. The power series provides the so-called *germ* of the analytical function, as it allows to calculate the function well away from the convergence radius of the series, by the well-known mechanism of analytic continuation.

As seen above, the algebraic problem has multiple solutions. Indeed at $s = 0$, multiple solutions as can be found for (3) if one allows $V_i(0) = 0$ and $\overline{V}_i(0) = 0$ at one or more buses i. However, those solutions are either *ghost* (unphysical) or *black* (non-vanishing injection at $s = 0$, meaning a short-circuit at the bus). On the other hand, by requiring $\overline{V}_i(0) \neq 0$ and $\overline{V}_i(0) \neq 0$ for all buses i, all injection terms in (3) vanish at $s = 0$ and the resulting linear system has a unique solution, dubbed the *white solution*. This is the operational solution of a power system under a no-load, no-generation scenario, where only the swing bus is providing a voltage source. The HELM method proposes that the operational solution to the load flow problem is the maximal analytical continuation of this reference solution at $s = 0$.

An interesting question may arise now, as to what solution is actually realized in a real power system. Is it necessarily the continuation of the white solution, as proposed by the method? To answer this question satisfactorily we need to remind ourselves that the load flow equations are just describing the *steady-state* of the physical system. To be certain about the actual state of the system we would need to integrate the differential equations of the full dynamical model, using some initial conditions. This would yield a unique steady state because there is no ambiguity in the solution of those equations. This is almost impossible to achieve in practice, but it is certain that the steady state will be analytic in the problem parameters, given some minimal assumptions of smoothness for the differential equations. In the absence of this modeling, solving the algebraic equations for the steady state actually requires to make a *proposal* as to what solution is the correct one for the physical systyem. In this sense, and invoking the smoothness of the underlying dynamical model, it seems most reasonable to propose the solution that is the analytical continuation of the white solution at the reference limit $s = 0$. Note however that the HELM procedure can also be applied to the calculation of black branches, and it possibly provides the best framework in which this can be done in a systematic manner.

The power series for the white solution can now be calculated as follows. Assume $V_i(s) = \sum_{n=0}^{\infty} c_i[n]s^n$ are the power series to be calculated, and $1/V_i(s) = \sum_{n=0}^{\infty} d_i[n]s^n$ are the corresponding power series for the functions appearing on the right hand side of (3). Making use of the reflection condition $\overline{V}_i(s) = V_i^*(s^*)$, the following equation is obtained where a formal power series appears on both sides:

$$\sum_k Y_{ik} \sum_{n=0}^{\infty} c_k[n]s^n = sI_i^{\text{load}} + sS_i^* \sum_{n=0}^{\infty} d_i^*[n]s^n \tag{6}$$

It is now straightforward to realize that (6) provides a way to progressively obtain the coefficients order by order, up to any desired level. The procedure is bootstrapped by making $s = 0$ in (6), which yields the linear system:

$$\sum_k Y_{ik} c_k[0] = 0$$

from which the zero-th order coefficients are obtained. Taking the derivative of (6) with respect to s and again evaluating at $s = 0$:

$$\sum_k Y_{ik} c_k[1] = I_i^{\text{load}} + S_i^* d_i^*[0] \tag{7}$$

But the coefficients $d_k[n]$ can be obtained from knowledge of the coefficients $c_k[m]$ (up to $m = n$) and $d_k[m]$ (up to $m = n-1$), since they are related by the convolution formulas:

$$\begin{aligned} 1 = V(s)V^{-1}(s) &= \left(\sum_{n=0}^{\infty} c[n]s^n\right)\left(\sum_{n=0}^{\infty} d[n]s^n\right) \\ &= c_0 d_0 + s\sum_{m=0}^{1} c[1-m]d[m] + s^2 \sum_{m=0}^{2} c[2-m]d[m] \\ &\quad + \cdots + s^n \sum_{m=0}^{n} c[n-m]d[m] + \ldots \end{aligned}$$

Thus $d_k[0] = 1/c_k[0]$ and the system (7) is readily solved to obtain the coefficients $c_k[1]$. The same procedure can now be repeated order after order:

$$\sum_k Y_{ik} c_k[n] = S_i^* d_i^*[n-1] \tag{8}$$

and by the convolution formulas:

$$d_i[n] = -\frac{\sum_{m=0}^{n-1} c_i[n-m]d_i[m]}{c_i[0]} \tag{9}$$

so that at every step a linear system is solved. Therefore the power series can be computed by a univocal and well-defined procedure which essentially consists in solving linear systems. In terms of computational work, it should be remarked that the matrix remains constant, and therefore its factorization only needs to be done once. Efficient implementations of the method should of course make use of modern sparse linear algebra routines [26], which include reordering algorithms for the minimization of the fill-up in the matrix factors. Section III-E analyzes the performance characteristics in further detail.

C. Analytic continuation

Since the procedure shown above can be carried out to arbitrary orders, a natural question arises as to how many orders it is needed to calculate, in order to obtain the final solution at the objective state $s = 1$. It should be first noted that the solution will not in general be obtained by direct summation of the power series at $s = 1$, as the radius of convergence is typically much smaller than 1. The powerful procedure of analytic continuation is used instead [27]. In passing, it should be stressed that the process of analytic continuation does not have anything in common with the concepts of numerical continuation (homotopy methods [19]) used in continuation load flow methods. In practice the analytical continuation is carried out by means of rational approximants, among which Padé approximation is the method of choice for reasons to be revealed shortly. As to the question of the number of terms needed in the power series in order to attain a given level of precision, practice has shown that typically anywhere from 10 to 40 terms suffice to reach 5-digit precision in large networks, and about 60 terms will exhaust the limits of the computer arithmetic in double precision.

The mechanics of the method have now been completely described, but another important issue needs to be addressed: is the analytical continuation procedure "complete"? That is, does it always reach the solution when it exists? Conversely, will the procedure unambiguously signal non-existence when the solution does not exist?. To answer these questions it is needed to invoke some powerful results from Complex Analysis.

As it is well-known from the theory of Algebraic Curves, all solutions $V_i(s)$ and $\overline{V}_i(s)$ are holomorphic functions that can be analytically continued along any path in s, as long as this path does not contain points of the exceptional set of the curve. This result can be alternatively stated by saying that Algebraic Curves are analytical functions whose only singularities are branch points, since the exceptional points of the curve (the points in s where the zeros of (5) have multiplicity greater than one, or equivalently, the points on the curve where $\partial\mathfrak{P}/\partial V_1 = 0$) are in fact branch points. These are the points where two or more branches of the curve $V_1(s)$ coalesce. The algebraic curve as a whole (all its branches) is a *complete global* analytic function [28], which means that there exist paths of analytic continuation connecting any points between any branches. In other words, knowledge of the power series at any (non-exceptional) point can be exploited to calculate the solution anywhere else, on any branch, by means of a suitable path of analytic continuation. However, note that analytical continuation paths enclosing branch points yield the curve on a branch different from the starting one. This is the subject of Monodromy Theory and is a foundational part of the theory of Riemann surfaces, which is the natural setting in which to study multivalued complex functions (and in particular, algebraic curves). For the purposes of the load flow method, the aim is to perform analytic continuation of the reference solution at $s = 0$ along paths that ensure single-valuedness, in other words, remaining always within the white branch. This is accomplished by the well-known procedure of selecting branch cuts on the complex plane. Branch cuts consist of lines connecting all branch points in such a way that no analytic continuation paths can encircle isolated branch points, in which case Monodromy Theory ensures that any chosen path for analytical continuation is single-valued.

However, since the specific geometry of cuts is arbitrary, some criteria need to be defined in order to choose them. At this point, two key results are invoked:

1) Stahl's extremal domain theorem [29]–[31]: this result asserts that for any analytic function there exists a unique set of cuts with the property of having minimal logarithmic capacity [32], and such that the function has single-valued analytic continuation in the domain consisting of the complex plane excluding the cuts (i.e., the maximal domain). This provides a natural criteria for the choice of cuts, and ensures that such choice exists and is unique.
2) Stahl's Padé convergence theorem [33], [34]: for any analytic function whose singularities are finite (in fact, it suffices for the set of singularities to have zero logarithmic capacity), any close-to-diagonal sequence of Padé approximants converge in capacity to said function in the extremal domain. The poles of the diagonal and paradiagonal Padé approximants accumulate on the set of cuts with minimal logarithmic capacity.

Padé approximants [31] are rational approximants to power series, and they have been used extensively as a technique for analytic continuation because their convergence has been known to be much better than that of power series. In particular, the diagonal and paradiagonal Padé approximants coincide with the continued fraction approximation to the power series, which are also known to have good convergence properties in general. Stahl's results, after the seminal works of Nuttall [35], reveal that Padé approximants are really a means for maximal analytic continuation. Therefore these two results confer the method very strong additional guarantees: if the Padé approximants converge at $s = 1$, the result is guaranteed to be the analytic continuation of the white branch at $s = 1$; conversely, if the Padé approximants do not converge at $s = 1$ (i.e. the point $s = 1$ lies on the set of cuts with minimal logarithmic capacity) then it is guaranteed that there is no solution (that is, the system is beyond voltage collapse).

D. The method in brief

To recapitulate, the Holomorphic Embedding Load Flow boils down to these steps:

1) Choose a suitable complex embedding by means of a complex parameter s. The embedding needs to be holomorphic (uses $V^*(s^*)$, not $V^*(s)$). At $s = 0$, this embedding should be such that the system becomes linear and trivially simple to solve (the no-load, no-generation case). This unambiguously selects the reference solution at $s = 0$.
2) Calculate the power series of $V(s)$ corresponding to the reference solution, by means of a sequence of linear systems that yield the coefficients progressively, order after order. The matrix in those systems remains always constant, so it needs to be factorized just once; and the right hand sides can always be calculated from the results of the previous system.
3) Compute the solution at $s = 1$ as the analytical continuation of the power series obtained in step 2, by using Padé Approximants. These are guaranteed to yield maximal analytical continuation, therefore the solution is obtained

when it exists, or a divergence is obtained when it does not exist.

Steps 2 and 3 are typically performed in an interleaved fashion, that is, after a new coefficient of the series is obtained, a new Padé approximant can be computed. The procedure stops when the desired accuracy is obtained, or when oscillation or divergence is detected. Section IV below will illustrate the whole procedure step by step.

E. Performance

The performance characteristics of the method are fairly easy to analyze in the asymptotic limit, when the size N of the network is large. Most of the work consists in factorizing a matrix system (8), and using the result to repeatedly solve different right-hand sides. Modern direct sparse solvers [26] are able to factorize a matrix A into its LDL^T Cholesky decomposition with a cost of $O(|L|)$, that is, linear in the number of non-zero entries obtained in the factorized matrix L. Therefore it is of utmost importance to use a fill-reducing reordering of the matrix. Solving each system adds $O(N)$ operations due to forward and back substitution. Computing the right hand sides of each linear system involves computing a simple convolution formula, reaching at most the order of the coefficient we are solving in the series. Since in practice there is a maximum of about 60 coefficients to solve before one encounters the limits of double-precision accuracy, the cost of calculation of these convolution formulas can considered to be fixed for each bus and thus contributes another $O(N)$. The same can be said about the computation of Padé Approximants. The cost of computing approximants increases approximately as the square of the number of coefficients being used, but since this number tops off at about a maximum of 60, the costs are still $O(N)$ asymptotically. Moreover, both the computation of the right-hand sides and the Padé approximants can be trivially parallellized. Therefore for large enough networks the costs are dominated by the factorization of the impedance matrix. These costs are then similar to the FDNR method, since the matrix factorization is only performed once.

To give some practical figures, in a real-world large transmission network of about 3,000 electrical nodes the HELM algorithm solves a load flow in about 10 to 20 ms, running on Intel Xeon 5500 CPUs, for a required precision of about 5 significant digits. The maximum orders for the Padé approximants are typically about 20, going up to a maximum of about 60 in scenarios where the network approaches voltage collapse.

IV. A WORKED EXAMPLE: THE TWO-BUS SYSTEM

The two-bus model offers an excelent playground to exemplify the HELM method and test the results against the exact known solution. Using the same sign convention as in (1), the equation for the system is:

$$\frac{V - V_0}{Z} = \frac{S^*}{V^*} \tag{10}$$

where V_0 represents the swing. In this case the linear algebra will be trivial since the system is one-dimensional. It is convenient to introduce adimensional variables by making $U \equiv \frac{V}{V_0}$ and $\sigma \equiv \frac{ZS^*}{|V_0|^2}$, so that the equation becomes

$$U = 1 + \frac{\sigma}{U^*} \tag{11}$$

which exhibits the essential algebraic structure of the load flow problem in its purest form. The exact solution is readily obtained by simple algebra manipulations:

$$U = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \sigma_R - \sigma_I^2} + j\sigma_I \tag{12}$$

subject to the condition $\Delta \equiv \frac{1}{4} + \sigma_R - \sigma_I^2 \geq 0$, where σ_R, σ_I are the real and imaginary parts of σ, respectively.

As summarized in Section III-D above, the HELM algorithm proceeds as follows. First, a suitable holomorphic embedding is proposed:

$$U(s) = 1 + \frac{s\sigma}{U^*(s^*)} \tag{13}$$

The second step consists in solving for the coefficients of the power series $U(s) = \sum c_n s^n$. For this we need the coefficients of the function $1/U(s) = \sum d_n s^n$. The system can be solved order after order, yielding the solution:

$$\begin{aligned} c_0 &= 1 \\ c_{n+1} &= \sigma d_n^* \qquad (n = 0, \ldots, \infty) \end{aligned} \tag{14}$$

where the coefficients d_n can always be obtained from those previously calculated, through the simple convolution formula $d_n = -\sum_{k=0}^{n-1} c_{n-k} d_k$. Here is the result for the first coefficients up to $n = 4$:

$$\begin{aligned} c_1 &= \sigma & c_3 &= \sigma^2\sigma^* + \sigma\sigma^{*2} \\ c_2 &= -\sigma\sigma^* & c_4 &= -\sigma^3\sigma^* - 3\sigma^2\sigma^{*2} - \sigma\sigma^{*3} \end{aligned}$$

The third step consists in the construction of the Padé Approximants using the power series coefficients. This is a standard procedure for which several efficient methods are available, such as Wynn's ϵ algorithm or the QD algorithm [31]. In this case it is instructive to use the general method, which involves solving a linear systems but provides the coefficients of the Padé approximants, instead of just its value at $s = 1$. Usually it is the diagonal and near-diagonal sequences of Padé approximants that give the best precision. Using the traditional $[L/M]$ notation, the first n coefficients of the series allow us to obtain approximants up to $L + M = n$:

$$\begin{aligned} [1/1] &= \frac{1 + (\sigma + \sigma^*)s}{1 + \sigma^* s} \\ [2/1] &= \frac{1 + (2\sigma + \sigma^*)s + \sigma^2 s^2}{1 + (\sigma + \sigma^*)s} \\ [2/2] &= \frac{1 + (2\sigma + 2\sigma^*)s + (\sigma^2 + \sigma\sigma^* + \sigma^{*2})s^2}{1 + (\sigma + 2\sigma^*)s + \sigma^{*2}s^2} \end{aligned}$$

It is now straightforward to evaluate these and compare them against the exact solution. Table I shows the precision obtained with some numerically computed Padé approximants, for three example scenarios. The third column corresponds to a case on the verge of collapse ($|V| = 0.58$).

It is worthwhile mentioning that the two-bus model explicitly exposes further insights into the problem. Looking at (13)

TABLE I
EXAMPLES OF RELATIVE PRECISION OF PADÉ APPROXIMANTS VS. ORDER

Padé order	$\sigma = -0.07 - j0.08$ ($\lvert V\rvert = 0.92$)	$-0.14 - j0.15$ ($\lvert V\rvert = 0.81$)	$-0.2 - j0.22$ ($\lvert V\rvert = 0.58$)
[2/2]	1.795e-03	2.30e-02	2.74e-01
[5/5]	7.02e-09	2.10e-05	5.85e-02
[10/10]	0	1.89e-10	1.16e-02
[15/15]	0	1.55e-15	2.86e-03
[20/20]	0	0	7.38e-04

and its complex comjugate, it is realized that the equations are telling us explicitely what the solution is, in *continued fraction* form. Just subsitute the denominators iteratively:

$$U(s) = 1 + \cfrac{\sigma s}{1 + \cfrac{\sigma^* s}{1 + \cfrac{\sigma s}{1 + \cfrac{\sigma^* s}{1 + \cdots}}}} \tag{15}$$

This is another representation of the function, but continued fractions have in general a much larger convergence radius and faster convergence compared to power series. In fact, it is well known that the convergents (i.e. truncations) of this continued fraction coincide with the diagonal and paradiagonal Padé approximants [31]. This may be verified against the approximants computed above in (15).

The general N-bus case does not easily lend itself into the elegant continued fraction approach, since there are linear systems involved. However, given the equivalence between Padé approximants and continued fractions, the analysis of the two-bus model is quite relevant, as the essential algebraic structure of the problem is already there.

V. FINAL REMARKS

A. Controls

Thus far only the pure PQ case has been considered. Real power systems have all sorts of automated controls imposing additional constraints on the solution of (1), the most pervasive being the PV controls for generator buses. These and other controls such as transformer ULTCs, FACTS, o HVDC links are easily accommodated under the HELM methodology. The explicit procedure for each type of constraint, although straightforward, is lengthy and will be the subject of a follow on paper [36]. Here it is pointed out that the only requirement is that the constraints can be expressed as algebraic equalities, which is true in all the aforementioned cases. Since no approximations are needed, the solutions benefit from all the mathematical guarantees that have been shown in this paper. This is in contrast to iterative methods, where controls are introduced as adjustments at each iteration. Such adjustments do not lend themselves to rigorous analysis and their effects on the solutions have to be studied empirically [37], [38].

B. Significance of HELM

Arguably the single most important impact of the HELM algorithm is the enabling of reliable, real-time, intelligent applications. The HELM method was actually born out of the need for a fully reliable load flow in the context of some AI-based applications that depend critically on the ability to perform exploratory load flow studies, with absolutely no margin for failure. The most prominent examples are two decision-support tools, a *Limits Violation Solver* and a *Restoration Plan Builder*. These tools are fully model-based thanks to a technique well-known in the AI community: guided exploration in the state-space of the electrical system, using the A* algorithm. The state-space consists of all possible electrical (steady) states that the network can achieve, and the available SCADA actions provide transitions between them. The algorithm needs sophisticated heuristics to guide the search efficiently, but the load flow method needs to be 100% reliable, as it is used at each and every step of the exploration. Our experience has shown that these kind of tools would be impossible to build on top of iterative load flow methods. Of course, other real-time tools such as Contingency Analysis or PV/QV Curves also benefit from increased reliability.

Another promising yet untapped potential of the method lies in the new insights it brings into the analysis of the load flow problem. The treatment in terms algebraic curves could prove quite powerful. For instance, it provides a coherent framework for the characterization and computation of all the multiple solutions to the original problem (white, black, ghost solutions). Given the vast amounts of results in the field of algebraic curves in Complex Analysis, it is reasonable to think that this is just scratching the surface of what is potentially possible. The theory of approximants (rational or other) is another source for insights and practical results. As it has been shown, the zeros and poles of the rational approximants tend to accumulate on the (minimal) branch cuts of the functions $V(s)$. Therefore their values, or even their patterns of appearance as the approximant order increases, may be used as new indicators, such as the proximity to voltage colapse. One may think of all this as a sort of a new language for the analysis of an old problem.

VI. CONCLUSIONS

This paper has presented a novel load flow method that radically breaks away from the established iterative methods. Its most salient features are that it is non-iterative, deterministic, and non-ambiguous: it guarantees, backed by mathematical proof, obtaining the right solution to the multivalued load flow problem, and otherwise signals unambiguously the non-existence of solution when the system is beyond voltage collapse. The method is based on a holomorphic embedding procedure that extends the voltage variables into analytic functions in the complex plane. This provides a framework to study and obtain the solutions using the full power of complex analysis. The method provides a procedure for constructing the complex power series at a well-defined reference point, where it is trivial to identify the correct branch of the multivalued problem, and then uses analytical continuation by means of algebraic approximants to reach the objective. It can be proven that the continuation is maximal in logarithmic capacity, thus propagating the chosen branch to the maximal possible domain on the complex plane. If the objective point is not in this domain, the initial correct branch does not have an analytical continuation there and the problem has no solution. Therefore, by construction, the process is totally deterministic thus ensuring that if there is a solution the method will find it and, conversely, if there is no solution (voltage collapse) the method will unequivocally signal such condition as well.

Although advanced results from geometric function theory are used to prove the method properties, the numerical implementation is straightforward and has performance characteristics that make it competitive and even superior to fast-decoupled algorithms, thus making it a good general-purpose load flow method for power systems of any size. The method has been implemented in industrial-strength EMS applications now operating at several large transmission operators, and has been granted two US Patents [39]. Experience has proven the foremost practical impact of the method, namely the enabling of *reliable* real-time applications.

ACKNOWLEDGMENT

The author would like to thank J. A. Marqués and V. Gaitán for their numerical implementations and extensive testing of the HELM algorithms since its inception, and J. L. Marín for his help in the preparation of this manuscript. The author also thanks the reviewers of the manuscript, whose comments helped improve the clarity of the paper.

REFERENCES

[1] A. Gómez-Expósito and F. L. Alvarado, "Load flow," in *Electric Energy Systems: Analysis and Operation*, A. Gómez-Expósito, A. J. Conejo, and C. Cañizares, Eds. CRC Press, 2009, pp. 95–126.

[2] J. B. Ward and H. W. Hale, "Digital computer solution of power-flow problems," *Power Apparatus and Systems, Part III. Transactions of the American Institute of Electrical Engineers*, vol. 75, no. 3, pp. 398–404, jan. 1956.

[3] W. Tinney and C. Hart, "Power flow solution by newton's method," *IEEE Trans. Power App. Syst.*, vol. PAS-86, no. 11, pp. 1449–1460, nov. 1967.

[4] W. Tinney and J. Walker, "Direct solutions of sparse network equations by optimally ordered triangular factorization," *Proc. IEEE*, vol. 55, no. 11, pp. 1801–1809, nov. 1967.

[5] S. Despotovic, B. Babic, and V. Mastilovic, "A rapid and reliable method for solving load flow problems," *IEEE Trans. Power App. Syst.*, vol. PAS-90, no. 1, pp. 123–130, jan. 1971.

[6] B. Stott, "Decoupled newton load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-91, no. 5, pp. 1955–1959, sept. 1972.

[7] B. Stott and O. Alsac, "Fast decoupled load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 859–869, may 1974.

[8] R. van Amerongen, "A general-purpose version of the fast decoupled load flow," *IEEE Trans. Power Syst.*, vol. 4, no. 2, pp. 760–770, may 1989.

[9] G. Leonidopoulos, "Approximate linear decoupled solution as the initial value of power system load flow," *Electric Power Systems Research*, vol. 32, no. 3, pp. 161–163, 1995.

[10] R. Klump and T. Overbye, "Techniques for improving power flow convergence," in *Power Engineering Society Summer Meeting, 2000. IEEE*, vol. 1, 2000, pp. 598–603.

[11] F. Wu, "Theoretical study of the convergence of the fast decoupled load flow," *IEEE Trans. Power App. Syst.*, vol. 96, no. 1, pp. 268–275, jan. 1977.

[12] J. Hubbard, D. Schleicher, and S. Sutherland, "How to find all roots of complex polynomials by newton's method," *Inventiones Mathematicae*, vol. 146, pp. 1–33, 2001.

[13] R. Klump and T. Overbye, "A new method for finding low-voltage power flow solutions," in *Power Engineering Society Summer Meeting, 2000. IEEE*, vol. 1, 2000, pp. 593–597.

[14] J. Thorp and S. Naqavi, "Load-flow fractals draw clues to erratic behaviour," *IEEE Comput. Appl. Power*, vol. 10, no. 1, pp. 59–62, jan 1997.

[15] Grupo AIA, "Problems with iterative load flow," http://www.elequant.com/products/agora/demo/iterativeloadflow, 2002–11.

[16] I. Dobson, "Observations on the geometry of saddle node bifurcation and voltage collapse in electrical power systems," *IEEE Trans. Circuits Syst. I*, vol. 39, no. 3, pp. 240–243, mar 1992.

[17] C. Cañizares, "On bifurcations, voltage collapse and load modeling," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 512–522, feb 1995.

[18] V. Ajjarapu and C. Christy, "The continuation power flow: a tool for steady state voltage stability analysis," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 416–423, feb 1992.

[19] E. Allgower and K. Georg, *Introduction to numerical continuation methods*, ser. Classics in applied mathematics. SIAM, 2003.

[20] A. Montes, "Algebraic solution of the load-flow problem for a 4-nodes electrical network," *Mathematics and Computers in Simulation*, vol. 45, no. 1-2, pp. 163–174, 1998.

[21] J. Ning, W. Gao, G. Radman, and J. Liu, "The application of the groebner basis technique in power flow study," in *North American Power Symposium (NAPS), 2009*, oct. 2009, pp. 1–7.

[22] W. Xu, Y. Liu, J. Salmon, T. Le, and G. Chang, "Series load flow: a novel noniterative load flow method," *Generation, Transmission and Distribution, IEE Proceedings*, vol. 145, no. 3, pp. 251–256, may 1998.

[23] A. Zambroni De Souza, C. Rosa, B. Lima Lopes, R. Leme, and O. Carpinteiro, "Non-iterative load-flow method as a tool for voltage stability studies," *Generation, Transmission Distribution, IET*, vol. 1, no. 3, pp. 499–505, may 2007.

[24] P. Sauer, "Explicit load flow series and functions," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 8, pp. 3754–3763, aug. 1981.

[25] B. Sturmfels, *Solving systems of polynomial equations*, ser. Regional conference series in mathematics. AMS, 2002.

[26] T. Davis, *Direct methods for sparse linear systems*, ser. Fundamentals of algorithms. SIAM, 2006.

[27] T. Gamelin, *Complex analysis*. Springer, 2001.

[28] L. Ahlfors, *Complex analysis: an introduction to the theory of analytic functions of one complex variable*. McGraw-Hill, 1979.

[29] H. Stahl, "Extremal domains associated with an analytic function I, II," *Complex Variables, Theory and Application: An International Journal*, vol. 4, no. 4, pp. 311–324, 325–338, 1985.

[30] ——, "The structure of extremal domains associated with an analytic function," *Complex Variables, Theory and Application: An International Journal*, vol. 4, no. 4, pp. 339–354, 1985.

[31] G. Baker and P. Graves-Morris, *Padé approximants*, ser. Encyclopedia of mathematics and its applications. Cambridge University Press, 1996.

[32] R. Rumely, *Capacity theory on algebraic curves*, ser. Lecture notes in mathematics. Springer-Verlag, 1989.

[33] H. Stahl, "On the convergence of generalized Padé approximants," *Constructive Approximation*, vol. 5, pp. 221–240, 1989.

[34] ——, "The convergence of Padé approximants to functions with branch points," *Journal of Approximation Theory*, vol. 91, no. 2, pp. 139–204, 1997.

[35] J. Nuttall, "On convergence of Padé approximants to functions with branch poits," in *Padé and Rational Approximation*, E. B. Saff and R. S. Varga, Eds. Academic Press, New York, 1977, pp. 101–109.

[36] A. Trias, "Holomorphic embedding load flow: Controls," 2012, preprint.

[37] S.-K. Chang and V. Brandwajn, "Solving the adjustment interactions in fast decoupled load flow," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 801–805, may 1991.

[38] J. Vlachogiannis, "Control adjustments in fast decoupled load flow," *Electric Power Systems Research*, vol. 31, no. 3, pp. 185–194, 1994.

[39] A. Trias, "System and method for monitoring and managing electrical power transmission and distribution networks," US Patents 7 519 506 and 7 979 239, 2009–2011.

Antonio Trias received a PhD degree in Physics from the University of Barcelona in 1974. From 1974 to 1976 he was a visiting postdoctoral research fellow at the Lawrence Berkeley Laboratory of the University of California. He is a founding member of Aplicaciones en Informática Avanzada S.A., where he currently is R&D Vice President and CTO. He has worked in the load flow problem and other algorithms of interest for power systems for over twenty years.