UC Davis
UC Davis Previously Published Works

Title

Computing resonant modes of accelerator cavities by solving nonlinear eigenvalue problems via rational approximation

Permalink

https://escholarship.org/uc/item/2rc9c3fs

Authors

Van Beeumen, Roel
Marques, Osni
Ng, Esmond G
et al.

Publication Date

DOI

Peer reviewed
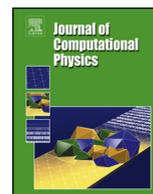
Computing resonant modes of accelerator cavities by solving nonlinear eigenvalue problems via rational approximation

Roel Van Beeumen [a],[*], Osni Marques [a], Esmond G. Ng [a], Chao Yang [a],
Zhaojun Bai [b], Lixin Ge [c], Oleksiy Kononenko [c], Zenghai Li [c], Cho-Kuen Ng [c],
Liling Xiao [c]

[a] *Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA, United States of America*
[b] *Department of Computer Science, University of California at Davis, Davis, CA, United States of America*
[c] *SLAC National Accelerator Laboratory, Menlo Park, CA, United States of America*

A R T I C L E   I N F O

A B S T R A C T

We present an efficient and reliable algorithm for solving a class of nonlinear eigenvalue problems arising from the modeling of particle accelerator cavities. The eigenvalue nonlinearity in these problems results from the use of waveguides to couple external power sources or to allow certain excited electromagnetic modes to exit the cavity. We use a rational approximation to reduce the nonlinear eigenvalue problem first to a rational eigenvalue problem. We then apply a special linearization procedure to turn the rational eigenvalue problem into a larger linear eigenvalue problem with the same eigenvalues, which can be solved by existing iterative methods. By using a compact scheme to represent both the linearized operator and the eigenvectors to be computed, we obtain a numerical method that only involves solving linear systems of equations of the same dimension as the original nonlinear eigenvalue problem. We refer to this method as a compact rational Krylov (CORK) method. We implemented the CORK method in the Omega3P module of the Advanced Computational Electromagnetic 3D Parallel (ACE3P) simulation suite and validated it by comparing the computed cavity resonant frequencies and damping Q factors of a small model problem to those obtained from a fitting procedure that uses frequency responses computed by another ACE3P module called S3P. We also used the CORK method to compute trapped modes damped in an ideal eight 9-cell SRF cavity cryomodule. This was the first time it was possible to compute these modes directly. The damping Q factors of the computed modes match well with those measured in experiments and the difference in resonant frequencies is within the range introduced by cavity imperfection. Therefore, the CORK method is an extremely valuable tool for computational cavity design.

Published by Elsevier Inc.

---

* Corresponding author.
E-mail addresses: rvanbeeumen@lbl.gov (R. Van Beeumen), oamarques@lbl.gov (O. Marques), egng@lbl.gov (E.G. Ng), cyang@lbl.gov (C. Yang), bai@cs.ucdavis.edu (Z. Bai), lge@slac.stanford.edu (L. Ge), oleksiy.kononenko@slac.stanford.edu (O. Kononenko), lizh@slac.stanford.edu (Z. Li), cho@slac.stanford.edu (C.-K. Ng), liling@slac.stanford.edu (L. Xiao).

1. Introduction

Particle accelerators are important experimental facilities for discoveries in physical sciences [9]. In the high energy frontier, particle colliders will be used to address many of the most compelling questions about dark matter, dark energy, and the fundamental nature of matter, energy, space and time. In basic energy sciences research, high-intensity particle beams or high-brightness light sources generated by particle accelerators will probe structures of matter that will benefit science discoveries in biology, chemistry, medical sciences, material sciences and others.

The design and optimization of accelerator cavities that are used to accelerate charged particles heavily depend on computational tools for achieving optimal design at reduced costs. With the advent of high performance computing (HPC), virtual prototyping of complex accelerator cavities has become a reality through high-fidelity simulation using advanced numerical algorithms. The parallel finite-element simulation suite ACE3P (Advanced Computational Electromagnetic 3D Parallel) [32], developed at the SLAC National Accelerator Laboratory, exploits both the power of HPC on state-of-the-art supercomputers and the accuracy of high-order finite elements for discretizing the partial difference equations that model the physical systems. ACE3P contains multiphysics modeling capabilities for integrated electromagnetic, thermal, and mechanical characteristics, and has been applied to a wide range of accelerator projects [18,13].

ACE3P's electromagnetic module Omega3P solves Maxwell's equations in the frequency domain [22]. Assuming the electromagnetic fields oscillating at a particular frequency, Maxwell's equations are mathematically cast into an eigenvalue problem, where the eigenvalue corresponds to the complex frequency of a resonant mode in an accelerator cavity, and the eigenvector represents the electric or magnetic fields of the resonant mode. In accelerator applications, a cavity is normally equipped with external waveguides either to couple power from external power sources into the cavity for acceleration or deflection of charged particles or to allow unwanted electromagnetic modes excited by the transit of a charged particle beam to exit the cavity. The eigenvalue problem is real, symmetric, and linear for a closed cavity without volumetric losses due to lossy materials inside the cavity, and without surface loss due to the finite electrical conductivity of the cavity walls, and correspondingly, the eigenvalues are real and represent the frequencies of the resonant modes.

In the presence of external coupling, the boundary conditions at the waveguide ports (or boundary surfaces) will render the discretized equation nonlinear with respect to the eigenvalue to be determined. Moreover, the eigenvalue λ is complex, with the real part related to the resonant mode frequency and the imaginary part related to the damping Q factor

$$Q = \frac{1}{2} \frac{\text{Re}(\lambda)}{\text{Im}(\lambda)}.$$

Typically in accelerator cavity design, the frequency ranges from MHz to GHz and the damping Q factor from a few tens to 10^7, depending on the modes of interest and the application. The eigenvalue problem yields a quadratic or a more nonlinear equation in the eigenvalue depending on the modes that can propagate through the external ports. When the waveguide modes propagate at the same propagation wave factor, the eigenvalue problem is quadratic and can be readily solved by standard methods [22]. When there are waveguide modes propagating at different propagation wave factors, the eigenvalue problem yields a more nonlinear equation in the eigenvalue. It is the purpose of this paper to describe an efficient algorithm for solving this type of nonlinear eigenvalue problem and show how such a solver can be implemented in Omega3P to address associated cavity design problems.

The paper is organized as follows. In section 2, we state the finite dimensional nonlinear eigenvalue problem in a more precise manner and review existing methods. In section 3, we describe our approach to solve the nonlinear eigenvalue problem, which is based on combining rational approximation and linearization techniques with a compact rational Krylov (CORK) subspace method. Next, we demonstrate the effectiveness of our approach in Section 4 with two numerical examples and compare the results to S3P simulations, which probe resonant responses to electromagnetic excitations by solving Helmholtz equations, and experimental measurements. Finally, the main conclusions are summarized in Section 5.

Throughout the paper, we denote scalars by lowercase Greek letters, vectors by lowercase Roman letters, matrices by capital Roman letters, and fields by bold capital Roman letters. The transpose and conjugate transpose of a matrix A are denoted by $A^\top$ and A^*, respectively. The matrix I_i is the identity matrix of dimensions i. We use V_j to denote a matrix with j columns, and its ith column is denoted by v_i. We omit subscripts when the dimensions of the matrices are clear from the context. The Kronecker product of the matrices A and B is denoted by $A \otimes B$.

2. Nonlinear eigenvalue problems in accelerator cavity simulations

2.1. Problem formulation

Maxwell's equations in differential form are

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \qquad \nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J},$$

$$\nabla \cdot \mathbf{D} = \rho, \qquad\qquad \nabla \cdot \mathbf{B} = 0,$$

where **E** is the electric field, **D** the electric displacement field, **B** the magnetic induction field, and **H** the magnetic intensity field, respectively. The **E**, **D**, **B**, and **H** fields satisfy the constitutive relations

$$\mathbf{D} = \varepsilon \mathbf{E}, \qquad \mathbf{B} = \mu \mathbf{H},$$

with ε and μ denoting, respectively, the permittivity and permeability of the medium. These parameters are tensors for anisotropic media and scalars for isotropic media. For inhomogeneous media, they are functions of the position, whereas for homogeneous media they are independent of the position. Both ε and μ are treated as constants in this paper.

When the field quantities are harmonically oscillating functions at a single frequency, Maxwell's equations can be expressed as

$$\nabla \times \mathbf{E} = -\mathrm{i}2\pi f \mathbf{B}, \qquad \nabla \times \mathbf{H} = \mathrm{i}2\pi f \mathbf{D} + \mathbf{J},$$

where f is the frequency. By eliminating **B**, **H**, and **D** or **E**, **B**, and **D**, we obtain

$$\nabla \times \left(\frac{1}{\mu} \nabla \times \mathbf{E} \right) - \varepsilon k^2 \mathbf{E} = 0, \tag{1}$$

$$\nabla \times \left(\frac{1}{\varepsilon} \nabla \times \mathbf{H} \right) - \mu k^2 \mathbf{H} = 0, \tag{2}$$

where $k = 2\pi f/c$ is the angular wavenumber and c the speed of light. Either equation can be used in the numerical simulation. Equation (1) is often known as the E-formulation and (2) as the H-formulation. Without loss of generality, we will use the E-formulation in this paper and solve this equation in Ω, which is the three-dimensional (3D) volume occupied by the cavity.

The boundary conditions for an electric field at a perfectly conducting surface Γ_E and for a magnetic field at a perfectly insulating surface Γ_M can be expressed as follows

$$\mathbf{n} \times \mathbf{E} = 0, \quad \text{on} \quad \Gamma_E, \tag{3}$$

$$\mathbf{n} \times (\nabla \times \mathbf{E}) = 0, \quad \text{on} \quad \Gamma_M, \tag{4}$$

where **n** is the surface normal. Note that in case the cavity exhibits symmetry, only a part of the geometry needs to be modeled. An accelerating cavity is often connected with waveguides to gain power or to damp high order modes. In order to terminate the propagation of the electromagnetic field, waveguides are truncated at a boundary port and an appropriate boundary condition is imposed at the port so that no reflection of the propagation field enters back into the computational domain. This is achieved by decomposing the field into waveguide eigenmodes, which should be distinguished from the cavity resonant eigenmodes, as [16]

$$
\begin{aligned}
\mathbf{n} \times \left(\frac{1}{\mu} \nabla \times \mathbf{E} \right) &- \mathrm{i}\gamma_0 e_0^{\text{TEM}} \int_\Gamma e_0^{\text{TEM}} \cdot \mathbf{E}\, \mathrm{d}\Gamma \\
&- \mathrm{i} \sum_m \gamma_m e_m^{\text{TE}} \int_{\Gamma_m} e_m^{\text{TE}} \cdot \mathbf{E}\, \mathrm{d}\Gamma_m \\
&+ \mathrm{i} \sum_m \frac{k^2}{\gamma_m} e_m^{\text{TM}} \int_{\Gamma_m} e_m^{\text{TM}} \cdot \mathbf{E}\, \mathrm{d}\Gamma_m = 0,
\end{aligned}
\tag{5}
$$

where $\gamma_0 = \sqrt{k}$ and $\gamma_m = \sqrt{k^2 - (k_m^c)^2}$, with $k_m^c = 2\pi f_m^c/c$ being the cutoff wavenumber of the mth waveguide mode, f_m^c being the corresponding cutoff frequency, and c being the speed of light. The normalized waveguide transverse electric and magnetic (TEM) mode, the mth transverse electric (TE) mode, and the mth transverse magnetic (TM) mode, onto the waveguide boundaries Γ_m, are given by e_0^{TEM}, e_m^{TE}, and e_m^{TM}, respectively.

We solve Maxwell's equations by the finite element method implemented in ACE3P. A high-order curved tetrahedral finite element mesh that conforms to the geometry of the structure is generated by Cubit [31]. Cubit uses ACIS as the basic geometry engine, which is used by many CAD software packages. The surfaces at the boundary of the computational domain are assumed to be piecewise smooth and represented as nonuniform rational B-spline (NURBS) surfaces. Nedelec edge elements [26] are employed to provide tangentially continuous basis functions for discretizing the electric field. The use of edge elements not only leads to a convenient way of imposing boundary conditions at material interfaces as well as at conducting surfaces, but also treats conducting and dielectric edges and corners correctly.

By representing the electric field in terms of hierarchical higher-order Nedelec basis functions $\mathbf{N}_i$ [35],

$$\mathbf{E} = \sum_i x_i \mathbf{N}_i,$$

where x_i is the ith component of the vector x, the E-formulation (1) together with the boundary conditions (3) to (5) lead to the following complex nonlinear eigenvalue problem

$$F(\lambda)x = 0, \tag{6}$$

where $\lambda = k^2$ is the eigenvalue, x the corresponding eigenvector, and

$$F(\lambda) = K - \lambda M + \mathrm{i}\sqrt{\lambda}\, W^{\mathrm{TEM}} + \mathrm{i}\sum_m \sqrt{\lambda - \kappa_m}\, W_m^{\mathrm{TE}} + \mathrm{i}\sum_m \frac{\lambda}{\sqrt{\lambda - \kappa_m}}\, W_m^{\mathrm{TM}} \tag{7}$$

a matrix-valued function, $F : \mathbb{C} \to \mathbb{C}^{n \times n}$, with $\kappa_m = (k_m^c)^2$ and

$$
\begin{aligned}
K &= \int_\Omega \frac{1}{\mu} \left(\nabla \times \mathbf{N}_i\right) \cdot \left(\nabla \times \mathbf{N}_j\right) \mathrm{d}\Omega, \\
M &= \int_\Omega \epsilon \left(\mathbf{N}_i \cdot \mathbf{N}_j\right) \mathrm{d}\Omega, \\
W^{\mathrm{TEM}} &= \int_\Gamma \left(e_0^{\mathrm{TEM}} \cdot \mathbf{N}_i\right) \cdot \left(e_0^{\mathrm{TEM}} \cdot \mathbf{N}_j\right) \mathrm{d}\Gamma, \\
W_m^{\mathrm{TE}} &= \int_{\Gamma_m} \left(e_m^{\mathrm{TE}} \cdot \mathbf{N}_i\right) \cdot \left(e_m^{\mathrm{TE}} \cdot \mathbf{N}_j\right) \mathrm{d}\Gamma_m, \\
W_m^{\mathrm{TM}} &= \int_{\Gamma_m} \left(e_m^{\mathrm{TM}} \cdot \mathbf{N}_i\right) \cdot \left(e_m^{\mathrm{TM}} \cdot \mathbf{N}_j\right) \mathrm{d}\Gamma_m.
\end{aligned}
\tag{8}
$$

Gaussian quadrature is used by ACE3P to evaluate the matrix elements of (8) numerically.

The eigenvalues to be computed depend on the resonant mode frequencies and damping Q factors of interest. Unlike linear or quadratic eigenvalue problems, this type of nonlinear eigenvalue problem is harder to solve. No standard software library is available for solving this type of problem, especially when the dimension of the problem is large.

2.2. Existing methods

The existing approaches to solve a general nonlinear eigenvalue problem can, roughly speaking, be grouped into three main classes [38]: Newton-based techniques, contour integration methods, and methods based on approximations of the nonlinear matrix-valued function.

The first approach treats the nonlinear eigenvalue problem as a system of nonlinear equations and tries to solve it by variants of Newton's method. See, e.g., [20,3,29,27] for some early references that tackle the nonlinear eigenvalue problem by solving a sequence of linear systems of equations. All these methods are local convergent methods that need a *good* initial guess. They are not guaranteed to converge. For large-scale problems, projection techniques are used to project the original problem into a properly constructed subspace, which is iteratively updated. Newton-like algorithms are used to solve the projected nonlinear eigenvalue problem. These methods include the nonlinear Jacobi–Davidson methods [7,10], the nonlinear Arnoldi [40] and nonlinear Rayleigh–Ritz [23] methods, and the block Newton method [19]. How to choose a *good* initial subspace remains an open question and the convergence of these methods can be very slow if no good starting guesses are available.

The second approach makes use of a contour integral reformulation of the nonlinear eigenvalue problem. In this approach, Keldysh's theorem [17] is used to convert the original nonlinear eigenvalue problem into a linear eigenvalue problem with identical eigenvalues inside a given contour. To obtain this linear eigenvalue problem, the contour integrals are computed via numerical quadrature and require the solution of a linear system with multiple right hand sides for every quadrature point. Examples are Beyn's method [8] and the nonlinear Sakurai–Sugiura method [4,44,25,41]. For a connection with rational filters, see, e.g., [37,36]. A drawback of all these methods is that the total number of matrix factorizations needed is the same as the number of quadrature points and, hence, these methods can become very computationally expensive.

The third approach, which we will use in this paper, is based on polynomial and rational approximation. Firstly, the nonlinear matrix-valued function $F(\lambda)$ is approximated by an interpolating polynomial or rational function. Next, the approximate polynomial or rational eigenvalue problem is solved by standard methods. For large-scale problems, several Krylov subspace methods have been developed to solve the linearizations of different approximate eigenvalue problems, such as the Chebyshev interpolation method [11], the rational linearization method [34], the infinite Arnoldi method [15], the Newton rational Krylov method [38], the fully rational Krylov method [14], and the generic class of CORK methods [39]. The key to make these methods efficient is to exploit the structure of the matrices resulting from linearization and the structure of the eigenvector. The reliability of all these methods heavily depends on the quality of the approximation of the nonlinear functions, but leads to more globally convergent methods than Newton-type methods. Moreover, its structure exploiting implementation in the CORK framework yields, in general, more efficient methods than using contour integration.

3. CORK nonlinear eigenvalue algorithm

The procedure we developed for solving the nonlinear eigenvalue problem (6)–(7) is composed of three steps: (i) *approximation* of the scalar nonlinear functions by interpolating rational functions, yielding a rational eigenvalue problem; (ii) *linearization* of the resulting rational eigenvalue problem, i.e., a reformulation of the rational eigenvalue problem as a generalized (linear) eigenvalue problem with the same eigenvalues but much larger in problem size; (iii) *solving* the generalized eigenvalue problem by a CORK method.

3.1. Approximation of square roots

The nonlinear eigenvalue problem (6) can, in contrast to linear eigenvalue problems, have an infinite number of eigenvalues and the corresponding eigenvectors can be linearly dependent. In practice, however, only a relative small number of eigenpairs are of interest. They typically lie in the so called *region of interest* $\Sigma \subset \mathbb{C}$. Therefore, we start by restricting the problem to finding only the eigenvalues λ inside this region of interest: $\lambda \in \Sigma$, and the corresponding eigenvectors.

Once the region of interest Σ is fixed, we approximate the matrix-valued function $F(\lambda)$ by a rational one $R(\lambda)$, so that the approximation error for all λ inside Σ is below a given tolerance $\varepsilon_{\text{approx}}$,

$$\max_{\lambda \in \Sigma} \|F(\lambda) - R(\lambda)\| \leq \varepsilon_{\text{approx}}.$$

It is important to select $\varepsilon_{\text{approx}}$ small enough since in [14] it has been illustrated that the accuracy of the approximate eigenvalues can be limited by the accuracy of the underlying (rational) approximation. Since $F(\lambda)$ is characterized by a summation of scalar nonlinear functions $f(\lambda)$ multiplied by matrices that do not depend on λ, we can take this specific structure into account when constructing a rational approximation $R(\lambda)$. In the remainder of the paper, we will make use of the rational Newton basis to approximate the square roots in (7).

We define the rational Newton polynomial of degree d, interpolating at the nodes $\sigma_0, \sigma_1, \ldots, \sigma_d$ and with nonzero poles $\xi_1, \ldots, \xi_d$, as follows

$$f(\lambda) \approx r(\lambda) := \sum_{j=0}^{d} \delta_j b_j(\lambda),$$

where δ_j are the *divided differences* defined as

$$\delta_0 := f[\sigma_0] = f(\sigma_0),$$
$$\delta_j := f[\sigma_0, \ldots, \sigma_j] = \frac{f[\sigma_1, \ldots, \sigma_j] - f[\sigma_0, \ldots, \sigma_{j-1}]}{\sigma_j - \sigma_0}, \qquad j = 1, \ldots, d,$$

and $b_j(\lambda)$ the *rational Newton basis functions*

$$b_0(\lambda) := 1,$$
$$b_j(\lambda) := \prod_{k=1}^{j} \frac{\lambda - \sigma_{k-1}}{\beta_k(1 - \lambda/\xi_k)}, \qquad j = 1, \ldots, d, \tag{9}$$

with $\beta_1, \ldots, \beta_d$ nonzero scaling parameters. Note also that they satisfy the following recurrence relation

$$b_j(\lambda) = \frac{\lambda - \sigma_{j-1}}{\beta_j(1 - \lambda/\xi_j)} b_{j-1}(\lambda), \qquad j = 1, \ldots, d.$$

The quality of the approximation heavily depends on the choice of the interpolation nodes and poles. In order to obtain a fast (low degree d) and quasi uniform approximation on the whole region of interest, we use near optimal point selection in Leja–Bagby fashion [6]. For more information see [14].

The rational approximation $R(\lambda)$ to (7) can efficiently be obtained by only approximating the (scalar) square root functions, yielding

$$F(\lambda) \approx R(\lambda) := \sum_{j=0}^{d} D_j b_j(\lambda), \tag{10}$$

where $b_j(\lambda)$ are the rational Newton basis functions defined in (9) and D_j the generalized divided difference matrices formed as linear combinations of the coefficient matrices of the original nonlinear matrix-valued function $F(\lambda)$. Finally, we obtain the rational eigenvalue problem

$$R(\lambda)x = 0, \tag{11}$$

which approximates the original eigenvalue problem (6) inside the whole region of interest up to the given tolerance $\varepsilon_{\text{approx}}$.

3.2. Linearization

Linearization is the classical approach for solving polynomial eigenvalue problems. Firstly, the matrix polynomial is transformed into a larger linear pencil with the same eigenvalues. Next, a linear eigensolver of choice can be applied to compute the eigenvalues of the polynomial eigenvalue problem. For polynomial eigenvalue problems, a companion matrix type of linearization is used in practice.

In order to linearize the rational eigenvalue problem (11), we make use of the linearization introduced in [14, Theorem 3.2]. Let $R(\lambda)$ be an $n \times n$ rational matrix-valued function as defined in (10). Then the $dn \times dn$ linear companion pencil

$$L(\lambda) = A - \lambda B,$$

where

$$A = \begin{bmatrix} D_0 & D_1 & \cdots & D_{d-2} & D_{d-1} - \sigma_{d-1} D_d/\beta_d \\ \sigma_0 I & \beta_1 I & & & \\ & \ddots & \ddots & & \\ & & \ddots & \beta_{d-2} I & \\ & & & \sigma_{d-2} I & \beta_{d-1} I \end{bmatrix}, \tag{12}$$

$$B = \begin{bmatrix} D_0/\xi_d & D_1/\xi_d & \cdots & D_{d-2}/\xi_d & D_{d-1}/\xi_d - D_d/\beta_d \\ I & \beta_1/\xi_1 I & & & \\ & \ddots & \ddots & & \\ & & \ddots & \beta_{d-2}/\xi_{d-2} I & \\ & & & I & \beta_{d-1}/\xi_{d-1} I \end{bmatrix}, \tag{13}$$

is a linearization of $R(\lambda)$. Furthermore, the eigenvectors of $L(\lambda)$ have a special structure: (i) if (λ, x) is an eigenpair of $L(\lambda)$, then there exists a vector x such that $x = b(\lambda) \otimes x$, with

$$b(\lambda) := \begin{bmatrix} 1 & b_1(\lambda) & \cdots & b_{d-1}(\lambda) \end{bmatrix}^\top, \tag{14}$$

and the pair (λ, x) is an eigenpair of $R(\lambda)$; (ii) if (λ, x) is an eigenpair of $R(\lambda)$, then the pair $(\lambda, b(\lambda) \otimes x)$ is an eigenpair of $L(\lambda)$.

3.3. CORK algorithm

After linearization, the nonlinear/rational eigenvalue problem is converted into a larger generalized eigenvalue problem. Note that by this operation the problem dimension has increased by a factor d, however, the resulting linear pencil has a particular structure which can be fully exploited if we compute the eigenvalues of interest via a Krylov method.

The standard rational Krylov method, which is generalization of the shift-and-invert Arnoldi method, is outlined in Algorithm 1. The method allows for varying the shift at every iteration. In every iteration we have the freedom to choose the next shift σ_j, e.g., based on the results of the previous iterations or we can cycle through some uniformly spaced shifts in order to obtain a faster and uniform convergence rate for all eigenvalues in the neighborhood of the chosen shifts.

Algorithm 1: Rational Krylov method [30].

Input: Matrices A and B.
Output: Approximate eigenpairs (λ_i, x_i).

1 Choose vector $v_1 \in \mathbb{C}^{dn}$, where $\|v_1\| = 1$.
  for $j = 1, 2, \ldots$ **do**
2   Choose shift σ_j and set continuation combination: t_j.
3   Shift-and-invert: $\hat{v} := (A - \sigma_j B)^{-1} B V_j t_j$.
4   Orthogonalize: $\tilde{v} = \hat{v} - V_j h_j$, where $h_j = V_j^* \hat{v}$.
5   Expand subspace: $V_{j+1} \leftarrow \begin{bmatrix} V_j & \tilde{v}/h_{j+1,j} \end{bmatrix}$, where $h_{j+1,j} = \|\tilde{v}\|$.
6   Compute eigenpairs: (λ_i, s_i) and test for convergence.
  end
7 Compute eigenvectors: $x_i = V_{j+1} \underline{H}_j s_i$.

The rational Krylov algorithm builds a subspace $V_j \in \mathbb{C}^{dn \times j}$ spanned by

$$v_1, \quad v_2 := (A - \sigma_1 B)^{-1} B v_1, \quad v_3 := (A - \sigma_2 B)^{-1} B v_2, \quad \cdots.$$

Eliminating $\hat{v}$ and $\tilde{v}$ in the jth iteration of Algorithm 1 and combining all the previous iterations yields the rational Krylov recurrence relation

$$AV_{j+1}\underline{H}_j = BV_{j+1}\underline{K}_j,$$

where $\underline{H}_j$ and $\underline{K}_j$ are two extended $(j+1) \times j$ upper Hessenberg matrices. Note that $\underline{H}_j$ and $\underline{K}_j$ denote the extended matrices formed by appending the extra row at the bottom of the Hessenberg matrices. The matrix $\underline{H}_j$ contains the coefficients of the Gram–Schmidt orthogonalization process and

$$\underline{K}_j = \underline{H}_j \operatorname{diag}(\sigma_1, \ldots, \sigma_j) + \underline{T}_j,$$

where the upper triangular matrix $\underline{T}_j \in \mathbb{C}^{(j+1) \times j}$ is built up from the continuation combinations $t_1, \ldots, t_j$, selected in the way described in [30].

Approximations for the desired eigenvalues and corresponding eigenvectors of the matrix pencil (A, B) can, in each iteration j of Algorithm 1, be obtained from the $j \times j$ upper parts of the two Hessenberg matrices $\underline{H}_j$ and $\underline{K}_j$

$$K_j s_i = \lambda_i H_j s_i, \qquad s_i \neq 0.$$

Then, we call $(\lambda_i, x_i := V_{j+1}\underline{H}_j s_i)$ a Ritz pair of (A, B).

The companion linearization matrices A and B, given in (12), (13), respectively, can be expressed in the following compact form

$$A = \left[\begin{array}{cccc} A_0 & A_1 & \cdots & A_{d-1} \\ \hline & M \otimes I_n & \end{array} \right], \quad B = \left[\begin{array}{cccc} B_0 & B_1 & \cdots & B_{d-1} \\ \hline & N \otimes I_n & \end{array} \right], \tag{15}$$

where $A_j, B_j \in \mathbb{C}^{n \times n}$ are related to the generalized divided difference matrices and $M, N \in \mathbb{C}^{(d-1) \times d}$ correspond to the linear relations between the rational Newton basis functions, i.e., $(M - \lambda N)b(\lambda) = 0$, for all λ, with $b(\lambda)$ defined in (14). Note that this compact structure of the linearization matrices (15) can be exploited by the CORK framework [39], which is a generic class of numerical methods for solving nonlinear eigenvalue problems.

The CORK method [39], which generalizes the two-level orthogonal Arnoldi procedure [24], is a special variant of the standard rational Krylov method and computes the eigenvalues and corresponding eigenvectors of the structured linearization pencil (15) by fully exploiting the Kronecker structure below the first block row, and by representing its subspace V_j in the following compact form

$$V_j = \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_d \end{bmatrix} = \begin{bmatrix} \mathcal{Q} & & & \\ & \mathcal{Q} & & \\ & & \ddots & \\ & & & \mathcal{Q} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_d \end{bmatrix} = (I_d \otimes \mathcal{Q})U,$$

where the matrix $\mathcal{Q} \in \mathbb{C}^{n \times r}$ is defined so that its columns form an orthonormal basis for the column space of the matrix $\begin{bmatrix} V_1 & V_2 & \cdots & V_d \end{bmatrix}$. In [39] it has been proven that the rank of $\mathcal{Q}$ is always bounded by $d + j$. Furthermore, this matrix $\mathcal{Q}$ will form an approximation to the corresponding eigenspace of the original nonlinear eigenvalue problem.

The particular structure of the linearization matrices (15) together with the compact representation of V_j in Algorithm 1 allows us to carry out steps 3–5 of Algorithm 1 very efficiently. This leads to the CORK algorithm, outlined in Algorithm 2, which only deals with matrices and vectors of the original nonlinear dimension n, instead of the linearization dimension dn in Algorithm 1.

Algorithm 2: Compact Rational Krylov (CORK) method [39].

Input: Matrices A_i, B_i, M, and N, defined in (15).
Output: Approximate eigenpairs (λ_i, x_i).

1　Choose vector $q_1 \in \mathbb{C}^n$, where $\|q_1\| = 1$.
　　for $j = 1, 2, \ldots$ **do**
2　　　Choose shift σ_j and continuation combination t_j.
3　　　Solve linear system: $F(\sigma_j)\hat{q} = y$.
4　　　Orthogonalize $\hat{q}$ against $\mathcal{Q}$.
5　　　Expand subspace: $\mathcal{Q} \leftarrow \begin{bmatrix} \mathcal{Q} & \hat{q}/\|\hat{q}\| \end{bmatrix}$ and update U.
6　　　Compute eigenpairs: (λ_i, s_i) and test for convergence.
　　end
7　Compute eigenvectors: $x_i = \mathcal{Q}U_1\underline{H}_j s_i$.

First, the shift-and-invert step in Algorithm 1, which involves a linear system solve of dimension dn, has been replaced by a linear system solve with the original nonlinear matrix-valued function (7) evaluated at the shift. The right hand side y in step 3 of Algorithm 2 can easily be computed as a sum of products of the nonlinear coefficient matrices (8) and linear combinations of the vectors in $\mathcal{Q}$. Second, the orthogonalization step in Algorithm 2 only deals with vectors of length n to update $\mathcal{Q}$ instead of vectors of length dn in Algorithm 1. Note also that the cost for updating U in step 5 of Algorithm 2 is negligible for large-scale problems. For more details on the implementation, we refer to [39].

Both the memory and computational costs in the CORK algorithm are significantly lower than the ones in the standard rational Krylov algorithm, so that for large-scale nonlinear eigenvalue problems the cost of the CORK algorithm is of the same order of magnitude as for just solving a generalized eigenvalue problem of the same dimension. Hence, the CORK algorithm allows us to efficiently and reliably compute the eigenvalues of interest of the nonlinear eigenvalue problem (6)–(8).

3.4. Implementation in Omega3P

Omega3P is the eigensolver module of ACE3P, which was originally designed for solving linear, generalized, and quadratic eigenvalue problems. We have implemented the CORK algorithm in Omega3P to solve nonlinear eigenvalue problems of the form (6)–(8) in order to also enable computing resonant modes of accelerator cavities equipped with waveguides.

After specifying the target region in which we want to compute eigenvalues, the square root functions in (7) are accurately approximated by rational Newton functions. We compute the corresponding generalized divided differences in a numerical stable way by using matrix functions [14]. This yields the actual rational eigenvalue problem to solve, where its coefficient matrices are formed as linear combinations of the original matrices K, M, and W's in (7).

The dominant cost of every CORK iteration is the linear system solve with $F(\sigma)$ in step 3 of Algorithm 2. For this operation, we make use of the sparse matrix computation infrastructure, such as sparse triangular factorizations, which was already in place in Omega3P.

4. Numerical results

In this section, we demonstrate the effectiveness of the CORK algorithm using two examples. In the first example, we compare the resonant frequencies and damping Q factors of a simple cavity model obtained from the numerical solution of the nonlinear eigenvalue problem (6)–(7) produced by the CORK algorithm to those obtained from ACE3P frequency-domain harmonic module S3P, which solves the Helmholtz representation of Maxwell's equations to determine the response to electromagnetic excitations at boundary ports of a cavity. In the second example, we use the CORK algorithm to validate for the first time ever the solution of a large-scale nonlinear eigenvalue problem against measurements obtained from a cryomodule consisting of eight superconducting cavities.

All numerical experiments were performed on the NERSC computers [21] Edison (Intel Xeon "Ivy Bridge" compute nodes @2.4 GHz, 2×12 cores, 64 GB DDR3 RAM) and Cori (Intel Xeon "Haswell" compute nodes @2.3 GHz, 2×16 cores, 128 GB DDR4 RAM). For solving the linear systems inside the nonlinear eigensolver, we used the software MUMPS, which is a multifrontal massively parallel sparse direct solver [1,2].

4.1. Pillbox cavity

We consider a simplified model of the RF gun cavity [43] used for the electron injector in the LCLS accelerator project to investigate the solution of the nonlinear eigenvalue problem obtained by the CORK algorithm. This simplified model is constructed by removing the cathode plate of the original LCLS RF gun cavity resulting in a pillbox cavity coupled with two rectangular waveguides and a circular beampipe of different radius at each end of the cavity. A number of TE and TM modes with different cutoff frequencies can propagate in the rectangular and circular waveguides at a particular frequency of the cavity resonant mode. Thus, the termination of the waveguide ports need to be loaded with the corresponding waveguide modes to ensure that the cavity mode couples to these waveguide modes without reflection at the ports. This results in a nonlinear eigenvalue problem as different nonlinear terms appear in the eigenvalue equation from these TE and TM modes loaded at the ports.

A quarter of the geometry of the pillbox cavity, as shown in Fig. 1, is used for the simulation by taking advantage of the two symmetry planes on which magnetic boundary conditions are set to reflect the model's symmetry. In this simulation, the targeted cavity resonant mode frequencies are below the lowest cutoff frequency ($\sim$2.2 GHz) of the smaller upstream circular beampipe. As a result, an electric boundary condition can be set on this beampipe end plane without affecting the eigenmode calculation as the fields of the resonant modes will be negligibly small there. Below the target frequency, there are three propagating modes in the rectangular waveguide and two in the bigger circular beampipe.

In both the Omega3P and S3P simulations, the rectangular ports are loaded with rectangular waveguide TE_{10}, TE_{11}, and TM_{11} modes with the corresponding cutoff frequencies $f_1^c = 0.908$ GHz and twice $f_4^c = 1.897$ GHz, respectively. The downstream circular ports are loaded with circular waveguide TM_{01} and TE_{21} modes with the corresponding cutoff frequencies $f_2^c = 1.043$ GHz and $f_3^c = 1.325$ GHz, respectively. These modes are required for the range of cavity mode frequency calculation as shown in Table 1. The first cavity mode can only couple to the rectangular TE_{10} and circular TM_{01} modes as its
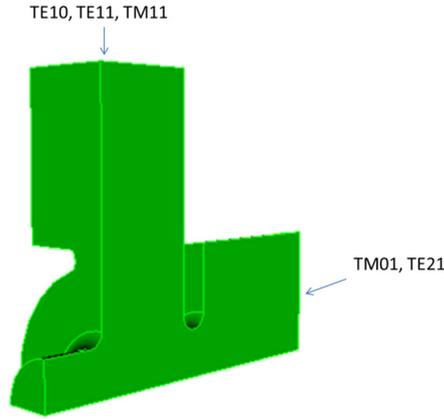
Fig. 1. A quarter model of a pillbox cavity coupled with rectangular waveguides and circular beampipes. The two symmetry planes are set to be magnetic boundary conditions. Three rectangular and two circular modes are loaded at the rectangular and downstream circular waveguides, respectively.

Table 1
Comparison of the cavity resonant mode RF parameters obtained by Omega3P (CORK algorithm) and indirect S3P calculations.

Mode	Omega3P (CORK)			S3P	
	f [GHz]	Q	$E(\lambda, x)$	f [GHz]	Q
1	1.1762	271	1.2827e−14	1.1762	270
2	2.0567	739	2.1098e−14	2.0567	760
3	2.1960	693	4.5892e−15	2.1965	780

frequency is above the cutoff frequencies of these two waveguide modes. The second and third cavity mode frequencies are above the cutoff frequencies of all the rectangular and circular waveguide modes mentioned above.

This configuration yields the nonlinear eigenvalue problem $F(\lambda)x = 0$, where

$$
F(\lambda) = K - \lambda M + \mathrm{i}\sqrt{\lambda - \kappa_1}\, W_{10}^{\mathrm{TE}} + \mathrm{i}\frac{\lambda}{\sqrt{\lambda - \kappa_2}} W_{01}^{\mathrm{TM}} \\
+ \mathrm{i}\sqrt{\lambda - \kappa_3}\, W_{21}^{\mathrm{TE}} + \mathrm{i}\sqrt{\lambda - \kappa_4}\, W_{11}^{\mathrm{TE}} + \mathrm{i}\frac{\lambda}{\sqrt{\lambda - \kappa_4}} W_{11}^{\mathrm{TM}},
\tag{16}
$$

and $K, M, W_{10}^{\mathrm{TE}}, W_{01}^{\mathrm{TM}}, W_{21}^{\mathrm{TE}}, W_{11}^{\mathrm{TE}}, W_{11}^{\mathrm{TM}}$ are symmetric matrices of dimension $n = 170,562$ (degrees of freedom) obtained by using an adaptive unstructured mesh with 27,384 quadratic tetrahedral elements and second order basis functions. Note that the nonlinearity in this nonlinear eigenvalue problem is due to the TE and TM waveguide mode propagation in the rectangular waveguides and the bigger downstream circular beampipe.

In the first calculation, we used the CORK nonlinear eigensolver to compute 3 cavity resonant modes inside the target frequency window [1.0 GHz, 2.2 GHz]. Note that this frequency window contains the cutoff frequencies f_2^c, f_3^c, and f_4^c. We have therefore subdivided it into 3 subwindows, which only contain 1 cutoff frequency at the time. We also applied in every subwindow a transformation of variables $\rho_i = \sqrt{\lambda - \kappa_i}$, for $i = 2, 3, 4$, to get rid of the branch point κ_i in the respective subwindow. Next, the transformed nonlinear functions were approximated by rational Newton functions of degree d at most 29, resulting in a relative approximation error smaller than 10^{-12} in every subwindow. For measuring the convergence of an approximate eigenpair (λ, x), we used the relative residual norm [23]

$$
E(\lambda, x) = \frac{\|F(\lambda)x\|_2/\|x\|_2}{\alpha(\lambda)},
$$

where $F(\lambda)$ is given by (16) and

$$
\alpha(\lambda) = \|K\|_1 + |\lambda|\|M\|_1 + \sqrt{|\lambda - \kappa_1|}\,\|W_{10}^{\mathrm{TE}}\|_1 \\
+ \frac{|\lambda|}{\sqrt{|\lambda - \kappa_2|}}\|W_{01}^{\mathrm{TM}}\|_1 + \sqrt{|\lambda - \kappa_3|}\,\|W_{21}^{\mathrm{TE}}\|_1 \\
+ \sqrt{|\lambda - \kappa_4|}\,\|W_{11}^{\mathrm{TE}}\|_1 + \frac{|\lambda|}{\sqrt{|\lambda - \kappa_4|}}\|W_{11}^{\mathrm{TM}}\|_1.
$$

The problem was solved within 3 minutes on the NERSC Cori computer [21] using 64 processors. The computed RF parameters (frequency f and damping Q factor) and relative residual norms are listed in Table 1. Note that all computed
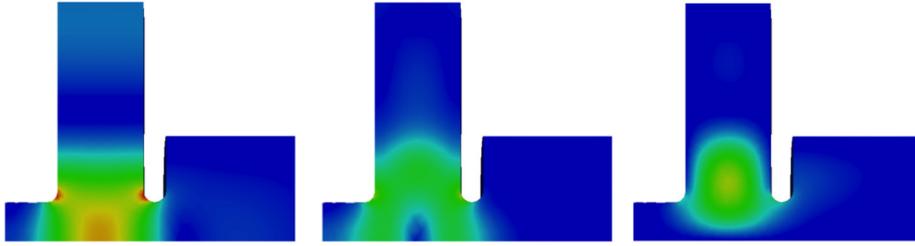
Fig. 2. Electric field amplitude profiles on the vertical symmetry plane for cavity modes at 1.176 GHz (left), 2.057 GHz (middle), and 2.196 GHz (right).
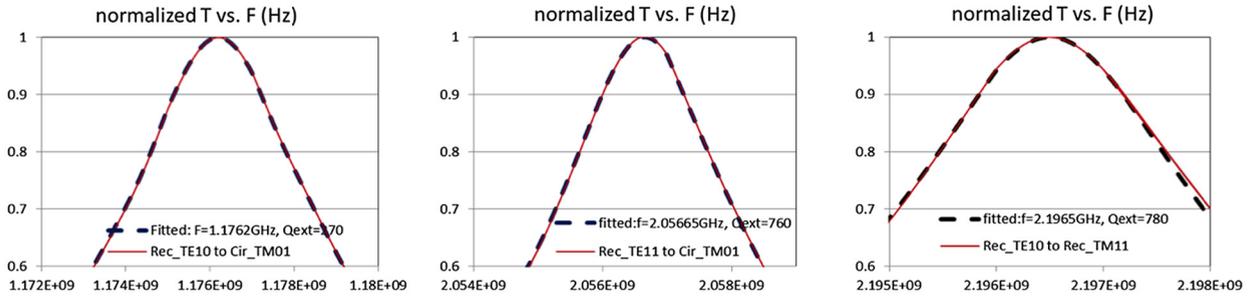
Fig. 3. Normalized transmission curves and the Lorentzian fit for the cavity modes at 1.176 GHz (left), 2.057 GHz (middle), and 2.196 GHz (right).

eigenvalues and corresponding eigenmodes are highly accurate since their relative residual norms are small. The corresponding electrical field patterns at a symmetry plane are shown in Fig. 2.

In the second calculation, we used the S3P approach by exciting waveguide modes at the waveguide ports at a certain frequency and determine the responses of the waveguide modes from the cavity. In particular, we can determine the scattering (S-) parameters (reflection and transmission coefficients) of each waveguide mode as a function of the frequency through a frequency scan around the resonant mode frequency to determine the damping Q factor from the width of an S-parameter curve. The resonant mode frequency f and damping Q can be obtained by fitting, for example, the transmission coefficient T with the Lorentzian as a function of frequency f' [28]

$$T^2 = \frac{1}{1 + Q^2(\frac{f'}{f} - \frac{f}{f'})^2}.$$

The corresponding results are given in the last 2 columns of Table 1. As illustrated in Fig. 3, the Lorentzian fits the normalized transmission curves from the S3P calculation for the three cavity modes with less than 1% fitting error within a 3 dB bandwidth.

In general, the cavity resonant mode results from the Omega3P–CORK eigensolver and S3P S-parameter solver agree well as shown in Table 1. The third cavity mode Q has more than 10% difference between the Omega3P and S3P results due to the contamination from nearby high order modes that causes the transmission curve asymmetry and thus a larger fitting error. In view of this, the CORK nonlinear eigensolver provides a direct way to compute the frequency and external damping factor of a cavity resonant mode without the need to perform a post processing fitting procedure from time-harmonic S-parameter calculations, which may be smeared by the overlap from other cavity modes in the spectrum.

4.2. TESLA SRF cavity cryomodule

It is also important to validate the CORK simulation results against experimental data to provide confidence for accelerator practitioners to use the method to design accelerator cavities through computation. Therefore, we now consider the cryomodule consisting of eight TESLA superconducting RF (SRF) cavities for which measurement data is available. The 9-cell TESLA SRF cavity, as shown in Fig. 4, has been adopted in many expensive accelerator projects, such as the proposed International Linear Collider (ILC) and the current Linac Coherent Light Source (LCLS-II) at SLAC. For example, 33 8-cavity cryomodules will be installed in the 4-GeV LCLS-II linear accelerator [33]. Hence, the accurate determination of the higher-order-mode (HOM) properties in the SRF cavities is very important for reliable operation of the machine.

The two HOM couplers on each side of the cavity, shown in Fig. 4, are designed to damp the dipole modes with two polarizations, while the fundamental coupler at the downstream is used to couple power into the cavity for the accelerating mode. In the simulation, the input coupler and the two HOM couplers are terminated by coaxial ports, where they can be modeled as absorbing boundary conditions in Omega3P. The frequencies of the lowest two dipole passband HOM modes are below the beampipe cutoff frequency at 2.253 GHz, and thus the two upstream and downstream beampipe ends can be
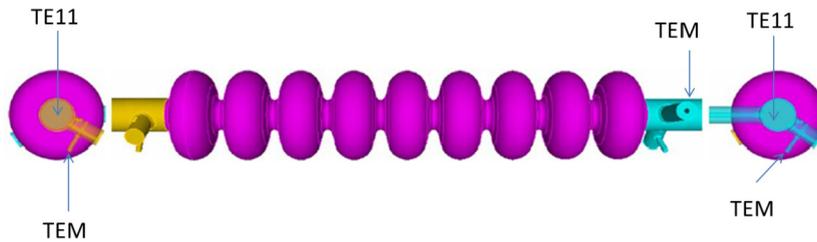
Fig. 4. TESLA cavity with one fundamental power coupler (FPC) and two higher-order-mode (HOM) couplers. The FPC and HOM couplers are loaded with a coax TEM mode and the two beampipe ends are loaded with two polarizations of the circular TE11 modes.
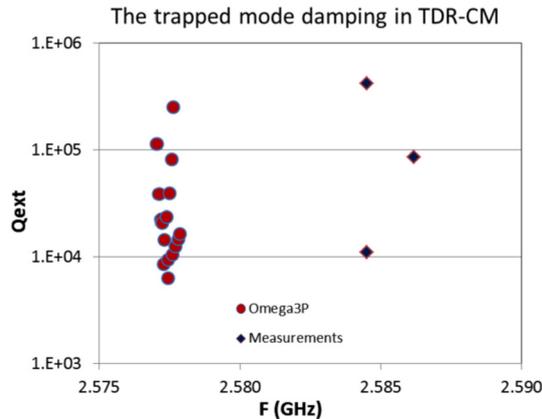
Fig. 5. Comparison of trapped mode damping calculations in an ideal LCLS-II 8-cavity cryomodule using the CORK nonlinear eigensolver in Omega3P and experimental measurements. The simulated frequencies are shifted for about 10 MHz to lower values due to cavity imperfection.

modeled as electric boundary conditions. In this case, the solution of the eigenmodes can be readily obtained by solving a quadratic eigenvalue problem. These two passband modes have been well studied through both numerical simulations and measurements [42].

The frequencies of the third dipole passband modes and beyond are above the beampipe cutoff frequency, and they can propagate through the cavities to the ends of each cryomodule, where absorbers are installed on the beampipe walls to damp these cavity modes. The last pair of the 3rd dipole band modes around 2.58 GHz has high transverse impedance R/Q, which may provide an unwanted transverse kick to the charged particle beam passing through the cavities. Therefore, adequate damping is required to suppress their potentially adverse effects to the beam. It was found from measurements that the dipole band modes were badly damped in one or two cavities in three TESLA cryomodules [5].

It is desirable to perform large-scale simulations of the cryomodule to corroborate with the experimental results. Since the 3rd dipole band mode frequencies are above the beampipe cutoff frequency at 2.253 GHz, the beampipe ends cannot be modeled as electric boundary conditions as for the two lowest dipole passbands, but rather, are terminated with waveguide boundary conditions, in which two polarizations of the circular TE_{11} waveguide modes need to be loaded. Together with the termination at the coupler ports as absorbing boundary conditions, simulating the 3rd dipole band trapped modes in the LCLS-II cryomodule yields a nonlinear eigenvalue problem.

Investigating the trapped modes in a full chain of accelerator cavities is computationally challenging. A qualitative eigenmode compendium method is developed to study the trapped modes in a 3.9 GHz European XFEL accelerator cryomodule, which is scaled down by a factor of 3 with respect to the TESLA cavity [12]. Prior to the implementation of the CORK algorithm, it was not possible to use existing tools in Omega3P to perform such a simulation. The addition of the CORK algorithm in Omega3P has enabled us to obtain the solution to the large-scale nonlinear eigenvalue problem resulting from a more realistic accelerator model directly.

In our simulation, a finite element mesh with 3 million quadratic tetrahedral elements is generated for the LCLS-II cryomodule consisting of 8 TESLA cavities. Using second order basis functions, we obtained a discretized problem with 20 million degrees of freedom. We used the CORK nonlinear eigensolver implemented in Omega3P to compute 16 trapped modes representing the two polarizations of the 8 coupled dipole modes inside the 8 cavities. The computation was completed within 10 wall clock minutes on the NERSC Edison computer [21] employing 960 processors. Fig. 5 shows the comparison of Omega3P calculations and experimental observations [5]. It can be seen that the calculated external Q values ranging from 10^4 to 10^6 of the 16 dipole modes covers very well the 3 measured values. It should be pointed out that the frequencies of the calculated modes deviate from the measured values. This is due to cavity imperfection arising from the deformation during the cavity fabrication process and the subsequent mechanical procedure to tune the cavity

Fig. 6. Electric field amplitude profile for the highest external Q mode in the 3rd dipole passband in an ideal LCLS-II 8-cavity cryomodule. The amplitude [a.u.] is represented by a linear rainbow color scheme. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

accelerating mode frequency to the design value. In doing this, the HOM dipole modes' frequencies will be shifted [42]. However, the damping of the HOM's will not be affected much as the mode's field patterns remain essentially the same, and the reliable determination of the Q values is crucial for assessing HOM effects on beam stability. To see this, we plot the electric field distribution of the highest external Q mode in Fig. 6. This plot shows that the intensity of the field at the beampipe ends becomes insignificant due to the high damping factor, which is consistent with experimental measurements.

5. Conclusions

We described the recently introduced CORK algorithm for solving a class of nonlinear eigenvalue problems arising from the modeling of particle accelerator cavities with waveguide couplings. The method relies on using a rational approximation to the nonlinear terms in the nonlinear eigenvalue problem to reduce the problem to a rational eigenvalue problem. A special linearization procedure turns the rational eigenvalue problem into a larger linear eigenvalue problem that can be solved by existing iterative methods. The efficiency of the method is achieved by using a compact scheme to represent both the linearized operator and the eigenvectors to be computed. This approach is more general and robust than algorithms that are based on using a fixed point iteration. We demonstrated the effectiveness of this method for both a model test problem and a more realistic TESLA SRF cavity cryomodule by implementing and using the CORK algorithm in Omega3P. The availability of this nonlinear eigensolver has enabled us to directly simulate high order modes of realistic cavity models with multiple waveguide coupling boundary conditions. The ability to perform such simulations efficiently and reliably will allow us to provide an accurate tool to design, optimize, and test sophisticated cavities ensuring the operational reliability of particle accelerators.

Acknowledgements

References

[1] P.R. Amestoy, I.S. Duff, J.-Y. L'Excellent, J. Koster, A fully asynchronous multifrontal solver using distributed dynamic scheduling, SIAM J. Matrix Anal. Appl. 23 (2001) 15–41, https://doi.org/10.1137/S0895479899358194.

[2] P.R. Amestoy, A. Guermouche, J.-Y. L'Excellent, S. Pralet, Hybrid scheduling for the parallel solution of linear systems, Parallel Comput. 32 (2006) 136–156, https://doi.org/10.1016/j.parco.2005.07.004.

[3] P.M. Anselone, L.B. Rall, The solution of characteristic value-vector problems by Newton's method, Numer. Math. 11 (1968) 38–45, https://doi.org/10.1007/BF02165469.

[4] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami, K. Kimura, A numerical method for nonlinear eigenvalue problems using contour integrals, JSIAM Lett. 1 (2009) 52–55, https://doi.org/10.14495/jsiaml.1.52.

[5] N. Baboi, M. Dohlus, C. Magne, A. Mosnier, O. Napoly, H.-W. Glock, Investigation of a high Q dipole mode at the TESLA cavities, in: Proceedings of European Particle Accelerator Conference, 2000.

[6] T. Bagby, On interpolation by rational functions, Duke Math. J. 36 (1969) 95–104, https://doi.org/10.1215/S0012-7094-69-03614-X.

[7] T. Betcke, H. Voss, A Jacobi–Davidson-type projection method for nonlinear eigenvalue problems, Future Gener. Comput. Syst. 20 (2004) 363–372, https://doi.org/10.1016/j.future.2003.07.003.

[8] W.-J. Beyn, An integral method for solving nonlinear eigenvalue problems, Linear Algebra Appl. 436 (2012) 3839–3863, https://doi.org/10.1016/j.laa.2011.03.030.

[9] A.W. Chao, H.O. Moser, Z. Zhao, Accelerator Physics, Technology and Applications, World Scientific, Singapore, 2004.

[10] C. Effenberger, Robust successive computation of eigenpairs for nonlinear eigenvalue problems, SIAM J. Matrix Anal. Appl. 34 (2013) 1231–1256, https://doi.org/10.1137/120885644.

[11] C. Effenberger, D. Kressner, Chebyshev interpolation for nonlinear eigenvalue problems, BIT Numer. Math. 52 (2012) 933–951, https://doi.org/10.1007/s10543-012-0381-5.

[12] T. Flisgen, J. Heller, T. Galek, L. Shi, N. Joshi, N. Baboi, R.M. Jones, U. van Rienen, Eigenmode compendium of the third harmonic module of the european x-ray free electron laser, Phys. Rev., Accel. Beams 20 (2017) 042002, https://doi.org/10.1103/PhysRevAccelBeams.20.042002.

[13] L. Ge, K. Ko, O. Kononenko, Z. Li, C.-K. Ng, L. Xiao, Advances in parallel finite element code suite ACE3P, in: Proceedings of International Particle Accelerator Conference, 2015.

[14] S. Güttel, R. Van Beeumen, K. Meerbergen, W. Michiels, NLEIGS: a class of fully rational Krylov methods for nonlinear eigenvalue problems, SIAM J. Sci. Comput. 36 (2014) A2842–A2864, https://doi.org/10.1137/130935045.

[15] E. Jarlebring, W. Michiels, K. Meerbergen, A linear eigenvalue algorithm for the nonlinear eigenvalue problem, Numer. Math. 122 (2012) 169–195, https://doi.org/10.1007/s00211-012-0453-0.
[16] J.-M. Jin, The Finite Element Method in Electromagnetics, 3rd ed., Wiley, 2014.
[17] M.V. Keldysh, On eigenvalues and eigenfunctions of some classes of non-self-adjoint equations, Dokl. Akad. Nauk SSSR 77 (1951) 11–14.
[18] K. Ko, A. Candel, L. Ge, A. Kabel, R. Lee, Z. Li, C. Ng, V. Rawat, G. Schussman, L. Xiao, Advances in parallel electromagnetic codes for accelerator science and development, in: Proceedings of LINAC10, 2010.
[19] D. Kressner, A block Newton method for nonlinear eigenvalue problems, Numer. Math. 114 (2009) 355–372, https://doi.org/10.1007/s00211-009-0259-x.
[20] P. Lancaster, A generalised Rayleigh quotient iteration for lambda-matrices, Arch. Ration. Mech. Anal. 8 (1961) 309–322, https://doi.org/10.1007/BF00277446.
[21] Lawrence Berkeley National Laboratory, National Energy Research Scientific Computing Center, 2017, http://www.nersc.gov/.
[22] L.-Q. Lee, Z. Li, C.-K. Ng, K. Ko, Omega3P: A Parallel Finite-Element Eigenmode Analysis Code for Accelerator Cavities, Technical Report SLAC-PUB-1352, SLAC National Accelerator Laboratory, 2009.
[23] B.-S. Liao, Z. Bai, L.-Q. Lee, K. Ko, Nonlinear Rayleigh–Ritz iterative method for solving large scale nonlinear eigenvalue problems, Taiwan. J. Math. 14 (2010) 869–883.
[24] D. Lu, Y. Su, Z. Bai, Stability analysis of the two-level orthogonal arnoldi procedure, SIAM J. Matrix Anal. Appl. 37 (2016) 195–214, https://doi.org/10.1137/151005142.
[25] N. Marsic, W. Ackermann, H. De Gersem, Non-linear eigenmode computations for conducting and superconducting cavities with a surface impedance boundary condition, IEEE Trans. Magn. 54 (3) (2018) 1–4, https://doi.org/10.1109/TMAG.2017.2742664.
[26] J.C. Nedelec, Mixed finite elements in $\mathbb{R}^3$, Numer. Math. 35 (1980) 315–341, https://doi.org/10.1007/BF01396415.
[27] A. Neumaier, Residual inverse iteration for the nonlinear eigenvalue problem, SIAM J. Numer. Anal. 22 (1985) 914–923, https://doi.org/10.1137/0722055.
[28] P.J. Petersan, S.M. Anlage, Measurement of resonant frequency and quality factor of microwave resonators: comparison of methods, J. Appl. Phys. 84 (1998) 3392–3402, https://doi.org/10.1063/1.368498.
[29] A. Ruhe, Algorithms for the nonlinear eigenvalue problem, SIAM J. Numer. Anal. 10 (1973) 674–689, https://doi.org/10.1137/0710059.
[30] A. Ruhe, Rational Krylov: A practical algorithm for large sparse nonsymmetric matrix pencils, SIAM J. Sci. Comput. 19 (1998) 1535–1551, https://doi.org/10.1137/S1064827595285597.
[31] Sandia National Laboratories, CUBIT Geometry and Mesh Generation Toolkit, http://cubit.sandia.gov/, 2017.
[32] SLAC National Accelerator Laboratory, ACE3P Simulation Suite, http://www-group.slac.stanford.edu/acd/, 2017.
[33] SLAC National Accelerator Laboratory, Linac Coherent Light Source II, https://portal.slac.stanford.edu/sites/lcls_public/lcls_ii/Pages/default.aspx, 2017.
[34] Y. Su, Z. Bai, Solving rational eigenvalue problems via linearization, SIAM J. Matrix Anal. Appl. 32 (2011) 201–216, https://doi.org/10.1137/090777542.
[35] D. Sun, J. Lee, Z. Cendes, Construction of nearly orthogonal Nedelec bases for rapid convergence with multilevel preconditioned Solvers, SIAM J. Sci. Comput. 23 (2001) 1053–1076, https://doi.org/10.1137/S1064827500367531.
[36] M. Van Barel, Designing rational filter functions for solving eigenvalue problems by contour integration, Linear Algebra Appl. 502 (2016) 346–365, https://doi.org/10.1016/j.laa.2015.05.029.
[37] M. Van Barel, P. Kravanja, Nonlinear eigenvalue problems and contour integrals, J. Comput. Appl. Math. 292 (2016) 526–540, https://doi.org/10.1016/j.cam.2015.07.012.
[38] R. Van Beeumen, K. Meerbergen, W. Michiels, A rational Krylov method based on Hermite interpolation for nonlinear eigenvalue problems, SIAM J. Sci. Comput. 35 (2013) A327–A350, https://doi.org/10.1137/120877556.
[39] R. Van Beeumen, K. Meerbergen, W. Michiels, Compact rational Krylov methods for nonlinear eigenvalue problems, SIAM J. Matrix Anal. Appl. 36 (2015) 820–838, https://doi.org/10.1137/140976698.
[40] H. Voss, An Arnoldi method for nonlinear eigenvalue problems, BIT Numer. Math. 44 (2004) 387–401, https://doi.org/10.1023/B:BITN.0000039424.56697.8b.
[41] J. Xiao, C. Zhang, T.-M. Huang, T. Sakurai, Solving large-scale nonlinear eigenvalue problems by rational interpolation and resolvent sampling based Rayleigh–Ritz method, Int. J. Numer. Methods Eng. 110 (2017) 776–800, https://doi.org/10.1002/nme.5441.
[42] L. Xiao, C. Adolphsen, V. Akcelik, A. Kabel, K. Ko, L. Lee, Z. Li, C. Ng, Modeling imperfection effects on dipole modes in TESLA cavity, in: Proceedings of Particle Accelerator Conference, 2007.
[43] L. Xiao, R.F. Boyce, D.H. Dowell, Z. Li, C. Limborg-Deprey, J.F. Schmerge, Dual feed RF gun design for LCLS, in: Proceedings of Particle Accelerator Conference, 2005.
[44] S. Yokota, T. Sakurai, A projection method for nonlinear eigenvalue problems using contour integrals, JSIAM Lett. 5 (2013) 41–44, https://doi.org/10.14495/jsiaml.5.41.