A generalized rational approximation of exponential integration (REXI) for massively parallel time integration

Martin Schreiber, Jed Brown

To cite this version:

Martin Schreiber, Jed Brown. A generalized rational approximation of exponential integration (REXI) for massively parallel time integration. 2023. hal-04363335

HAL Id: hal-04363335
https://hal.science/hal-04363335
Preprint submitted on 24 Dec 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License
A GENERALIZED RATIONAL APPROXIMATION OF
EXPONENTIAL INTEGRATION (REXI) FOR MASSIVELY
PARALLEL TIME INTEGRATION

MARTIN SCHREIBER∗ AND JED BROWN†

Abstract. Solving partial differential equations (PDEs) is one of the most traditional tasks in
scientific computing. In this work, we consider numerical solutions of initial value problems (IVPs)
problems partly or entirely given by linear PDEs and how to compute solutions with a method
we refer to as rational approximation of exponential integration (REXI). REXI replaces a typically
sequential timestepping method with a sum of rational terms, leading to the possibility to parallelize
over this sum. Hence, this method can potentially exploit additional degrees of parallelization for
scaling problems limited in their spatial scalability to large-scale supercomputers.

The main contribution of this work lies in developing the “unified REXI” in which we show
algebraic equivalence to other methods developed up to five decades ago. Such methods cover, e.g.,
diagonalization of the Butcher table for implicit Runge-Kutta methods, Cauchy-contour integration-
based methods, and direct approximations. To our best knowledge, this is the first time of such a
comparison and deep investigation of all these methods.

Finally, we will show the applicability of REXI to the nonlinear shallow-water equations on the
rotating sphere, including HPC results. While previous REXI studies have focused on exposing more
parallelism to enable faster time to solution, we also consider efficiency at prescribed accuracy and
find that diagonalized Gauss Runge-Kutta methods (formulated as REXI) are compelling highly
efficient methods.

Key words. Exponential integrators, rational approximation, parallel-in-time, Cauchy contour,
Butcher table, diagonalization

AMS subject classifications.

1. Introduction. Time integration of IVPs is one of the most traditional tasks in
scientific computing, having seen two centuries of research. The IVPs we are interested
in are given entirely or partly by linear autonomous PDEs, which are ubiquitous in
applications ranging from daily weather forecasting [11] to full waveform inversion
[43]. Integration of such systems is sequential in time using conventional methods such
as explicit and diagonally implicit Runge-Kutta [29, 21]: Without special structure
[20], the state at each stage is necessary to compute the next stage, either explicitly
or implicitly. The time step size is typically limited by stability and/or accuracy
requirements and the method is purely sequential in the time dimension.

With the desire to solve PDEs with ever-higher resolutions, the demands on high-
performance computers (HPC) have increased. The steady and ongoing increase in

∗Univ. Grenoble Alpes / Laboratoire Jean Kuntzmann / Inria, Grenoble, France (martin.schreiber@univ-grenoble-alpes.fr), Department of Informatics, Technical University of Munich, Germany (martin.schreiber@tum.de), https://www.martin-schreiber.info
†Department of Computer Science, University of Colorado, Boulder, USA (jed@jedbrown.org, https://jedbrown.org)

Submitted to the editors DATE.

Funding: This project has received funding from the Federal Ministry of Education and Re-
search and the European High-Performance Computing Joint Undertaking (JU) under grant agree-
ment No 955701. The JU receives support from the European Union’s Horizon 2020 research and
innovation programme and Belgium, France, Germany, Switzerland.

Martin Schreiber gratefully acknowledges KONWIHR funding as part of the project “Parallel in
Time Integration with Rational Approximations targeting Weather and Climate Simulations”.

Jed Brown acknowledges support from the U.S. Department of Energy, Office of Science, Office of
Advanced Scientific Computing Research, applied mathematics program.
HPC performance is provided almost exclusively by increased parallelism; increasing resolution in space (spatial scalability) can be solved in the same amount of time per time step, but the wallclock time to simulate for a fixed physical duration increases due to the increasing number of time steps to satisfy the Courant-Friedrichs-Lewy (CFL) constraint \[9\] for transport phenomena. Consequently, refinement to increase accuracy on a transient physical problem is always a scaling challenge, and many applications are unable to increase spatial resolution without sacrificing external timelines such as IPCC assessment reports \[6\] or design/manufacturing timelines. Parallelism in the time dimension is seen as an opportunity to utilize greater parallelism to meet stringent simulation timelines. The Rational Approximation of Exponential Integration (REXI) family of methods, which we briefly explain next, are a promising candidate for hyperbolic PDEs. Consider a linear autonomous PDE given by
\[
\frac{\partial U(t)}{\partial t} = LU(t)
\]
with \(L\) the discrete linear operator and \(U(t)\) the discrete state variables at time \(t\). Solving such IVPs have been intensively studied over the last decades with various approaches, and one of the direct methods is the application of an exponential integration
\[
U(t + \Delta t) = \exp(\Delta tL)U(t)
\]
with the solution \(U(t)\) at time \(t\). We want to emphasize that no time discretization has been introduced and that the only approximations are related to space. The REXI method exploits the feature that \(\exp(\Delta tL)\) only needs to be approximated within a spectrum related to time step size \(\Delta t\) and the spectrum of \(L\). In the present work, we can express a variety of different time integration methods by what we refer to as the “unified REXI” formulation given by
\[
U(t + \Delta t) \approx \gamma U(t) + \sum_{n=1}^{N} \beta_n (\Delta tL - \alpha_n)^{-1} U(t)
\]
with the time step size given by \(\Delta t\), and (typically) complex valued REXI coefficients \(\alpha_n, \beta_n\) and real-valued \(\gamma\) is a new generalization that we will use in the following sections. The remainder of this work investigates different ways to infer these REXI coefficients and their relation to a class of Runge-Kutta methods. Based on this, we will study their numerical properties in the linear and nonlinear context.

2. Related Work.

2.1. Exponential integration. Exponential integration methods are formulated for nonlinear systems factored as
\[
\frac{\partial U(t)}{\partial t} = LU(t) + N(U(t))
\]
where the linear part \(L\) is intended to capture the “fast” dynamics and \(N\) is the remaining nonlinear part. An exact ansatz for advancing this split equation over a

This manuscript is for review purposes only.
A GENERALIZED REXI FOR MASSIVELY PARALLEL TIME INTEGRATION

A finite time interval is given by

\[
U(t + \Delta t) = \exp(\Delta tL)U(t) + \int_0^{\Delta t} \exp((\Delta t - \tau)L)N(U(t + \tau))d\tau.
\]

In this form, the linear parts are integrated precisely by an exponential function, hence overcoming potential stiffness challenges caused by the linear parts. Due to this advantageous property, the interest in these exponential integrators has steadily increased over the last decades (see, e.g., [25, 19]) where various approaches have been taken to approximate the integral of the nonlinearities. One of the most commonly known approximations of the integral is, e.g., given by (see [10])

\[
U(t + \Delta t) = \varphi_0(\Delta tL)U(t) + \Delta t\varphi_1(\Delta tL)N(U(t))
\]

where we used the notations \(\varphi_0(Z) = e^Z\) and \(\varphi_1(Z) = \frac{e^Z - 1}{Z}\). We skip further examples for discretized exponential integrator formulations and only like to point out the \(\varphi\) functions to be omnipresent in higher-order exponential integration methods, which is generally given, e.g., by

\[
\varphi_{i+1}(Z) = (\varphi_i(Z) - \varphi_i(0))Z^{-1} \quad \text{for } i \geq 0.
\]

An investigation of all different varieties of discretizations of exponential integrators incorporating the nonlinearities is beyond the scope of this work, and we continue on the linear parts. These linear parts can be either given by full linear PDEs or by time integrating only a part of linear PDEs where the underlying requirement of time integration results in problems of the form \(U(t + \Delta t) = \varphi_0(\Delta tL)U(t) = \exp(\Delta tL)U(t)\). In contrast to state-of-the-art time integration methods, which are used in operational codes, exponential integrators for linear operators avoid any time-discretization errors. However, the computational complexity can be tremendous and triggered the development of various ways to tackle this challenge [25]. We will briefly summarize the ones recently researched, namely based on Krylov subspaces and REXI.

The exponential can be approximated using Krylov subspace solvers (see [26, 39, 40, 8]) where we see polynomial approximations (e.g., based on Chebyshev) as a subclass of them. The advantage of such methods is their simplicity – assuming the Krylov solver framework given – since only vector multiplications with the linear operator are required. However, the potential drawbacks of Krylov subspace solvers are their inherent property of sequential iterations over the Krylov subspace, hence not providing ways to exploit additional degrees of parallelization. An alternative is to use the REXI method, which will be discussed in the next section in further detail.

### 2.2. REXI

The particular way to evaluate the \(\varphi\) functions in the present work is strongly related to Padé approximations, which can be used as a first instance to approximate the \(\varphi\) functions. This approximation is most naturally related to how all Runge-Kutta formulations, e.g., based on the Butcher table, can be formulated for linear autonomous operators. However, higher-order polynomials in the denominator also make the development of solvers for such Padé approximations more challenging, and a partial fraction decomposition can be used. This well-known decomposition transforms a higher-order Padé approximation into a sum of lower-order terms, which can eventually be used to develop solvers parallelized over all terms. Although not being explained in the context of Padé approximations, REXI methods using the partial fraction decomposed form have been developed in different contexts. They
can be interpreted as a Padé approximation, which is why they are mentioned in this context.

In what follows, we will provide an overview of different methods, which can all be phrased in REXI form. One of the earliest REXI formulations for hyperbolic PDE time integrators is related to the Laplace transformation (cf. [23, 7]). Here, the PDE is transformed with the Laplace operator, where the backward transform is conducted with a Cauchy Contour integral. This transformation can be again related to an exponential integration scheme, namely to the Cauchy Contour method mentioned below, see also [41]. More recently, time integration based on Laplace transformations with a circle-based Cauchy contour integration have been more intensively studied in [28] with ODEs, in particular filtering properties. However, it needed a more extensive (community) effort to develop other, e.g., higher-order methods around them, as has been extensively the case for exponential integration methods. This lack of advancements with the REXI Laplace transform is also why we only concentrate on exponential integration-based formulations. We want to point out that the same approaches could also be taken from the Laplace transform perspective.

Another way to infer REXI coefficients originates from the REXI method based on Gaussian basis functions originally developed in [18], which only targets purely oscillatory problems (hence $L$ has only imaginary or zero eigenvalues). This method also showed excellent properties regarding the wallclock-time vs. error for the linear shallow-water equations on the plane (see [33]) and on the rotating sphere [32, 34].

Although initially developed for analytical reasons, the Cauchy contour integration method can indeed be used for REXI time integration. As pointed out above, one of the first times this has been used as a REXI-like method was with the Laplace transformations. However, exponential formulations (see Eq. 2.2) provide a more direct and substantial established way to integrate in time. Here, the property of $\phi_i$ being an analytical function plays a fundamental role in the Cauchy contour integration method as well as the exponentially fast converging trapezoidal rule to approximate the contour [41]. This method has already been used in different works: The approximation of $\phi_i(x)$ evaluations on scalar values has been used in various works to overcome singularities of $\phi_i>0$ singularities at the origin, see, e.g., [5]. It has been used mainly for parabolic problems [35], also pointing out the potential of parallelization for the first time, as well as using a Carathéodory-Fejér method [31]. Regarding real applications, it was applied to nonlinear shallow-water equations on the rotating sphere [34], providing improved wallclock time-to-solution by using an enlarged and shifted contour to avoid numerical cancellation errors.

### 2.3. Parallel-in-time

Overcoming the wallclock time limitations of simulations, which cannot be accomplished by any further increase of parallelization in the spatial dimension, is the main focus of the parallel-in-time algorithms. Here, two different types of approaches exist: (a) minimally-invasive methods that take existing time integration methods and incorporate them into an iterative-in-time correction scheme (see, e.g., Parareal [22] and PFASST [24]); and (b) invasive methods that replace an existing time stepping with one that works entirely differently. Very often, one likes to use a combination of these approaches to enhance the convergence speed of the correction scheme in time. REXI is an invasive parallel-in-time algorithm (see [33]) since it requires efficient complex-valued solvers for each REXI term.

### 3. Unified REXI formulation

We start directly with the REXI formulation which will provide a standard fundament for the different variants to infer REXI coefficients. Given a discrete linear operator $L$, we can use an eigendecomposition
A GENERALIZED REXI FOR MASSIVELY PARALLEL TIME INTEGRATION

$L = QΛQ^{-1}$ with the eigenvectors stored in the columns of $Q$ and the eigenvalues placed correspondingly on the diagonal of $Λ$.

\[
\frac{∂U(t)}{∂t} = LU(t) = QΛQ^{-1}U(t)
\]

where $Q$ and $Λ$ are the matrices with the eigenvectors and eigenvalues on the diagonal, respectively. In terms of the characteristic variable $u = Q^{-1}U$ and due to diagonal-$Λ$, we get independent equations of the form

\[
\frac{∂u_i(t)}{∂t} = λ_i u_i(t)
\]

with $λ_i$ the individual Eigenvalues on the diagonal of $Λ$. In characteristic variables, the unified REXI formulation (1.3) becomes

\[
u_i(t + Δt) ≈ γu_i(t) + \sum_{n=1}^{N} β_n(Δtλ_i - Iα_n)^{-1}u_i(t).
\]

Since each component $u_i$ is decoupled, we can freely drop the subscript. For the purpose of time integration, the linear operator $L$ is completely described by its eigenvalues $λ$, where imaginary components $ℑ(λ)$ represent oscillation and negative real values $ℜ(λ) < 0$ describe a diffusive/damping behavior. Note that substituting $λ = 1, t = 0, Δt = x, u(0) = 1$ in (3.2) yields $exp(x) = γ + \sum β_n(x - α_n)^{-1}$, which provides intuition as a sum of rational functions.

3.1. Exploiting symmetry of coefficients. We note that it is possible to reduce the workload by a factor of two for real-valued operators $L$ when the poles $α$ consist of complex conjugate pairs (see, e.g., [23, 18]). This optimization does not change the relative performance of the methods we consider here, so for simplicity, we do not apply it.

3.2. REXI-derived higher-order $ϕ$ forms. Particular higher-order exponential time integrators such as (2.3) require evaluations of higher-order $ϕ_{i|i>0}$. REXI coefficients for these functions are so far computed with methods tailored to them, see [18, 32]. We briefly present an new alternative way to compute them which is easily applicable. Given REXI coefficients for

$ϕ_i(x) ≈ γ + \sum β_n(x - α_n)^{-1}$

we can compute higher-order REXI approximations with

$ϕ_{i+1}(x) = \frac{ϕ_i(x) - ϕ_i(0)}{x} = \frac{γ + \sum n β_n(x - α_n)^{-1} - ϕ_i(0)}{x}$

$= \sum \frac{β_n}{α_n(x - α_n)} + \frac{1}{x} \left( \sum \frac{β_n}{α_n} + γ - ϕ_i(0) \right) = \sum \frac{β_n}{x - α_n}.$

The cancellation of the terms is a consequence of the stationary modes which require $\sum_n \left( \frac{β_n}{α_n} \right) + γ = ϕ_i(0)$. Note that this leads to different coefficients compared to tailored computations.
3.3. Linear solvers for REXI terms. Efficient solvers are required for each REXI term. Over the last decades, this efficiency aspect turned out to be a very challenging task. E.g., in the context of shallow-water equations, this results in the original Helmholtz problem (rather than a backward Euler time step) where it is known that no off-the-shelf solvers such as GMRES and multigrid methods work in a highly-scalable way (see, e.g., [12]). This is ongoing research, and in the present work, we are using solvers developed in spherical harmonics formulations, hence a solver tailored particularly to one PDE problem.

4. REXI methods. These sections cover various ways to infer REXI coefficients, which represent, from our point of view, the most interesting cases. The goal is not to show all methods in great detail but their fundamental properties.

Although we present methods in characteristic form (3.2), the proposed methods also hold in system form (1.3). In the following, we will use the error

\[ e(z) = \gamma + \sum_n \beta_n (z - \alpha_n)^{-1} - \exp(z) \]  

to compute the deviation from \( \varphi_0(z) = \exp(z) \) with \( z = \lambda \Delta t \) denoting the point on the complex plane to evaluate. Since approximating diffusive problems is relatively straightforward, we focus on purely oscillatory problems with \( \lambda \in i \mathbb{R} \). The REXI methods we consider have complex-conjugate poles \( \alpha \), thus \( e(z) = e(\bar{z}) \) and so we only plot errors for \( \Im[z] \geq 0 \).

4.1. B-REXI: Butcher/Bickart. A Butcher table [2] provides a canonical representation of s-stage Runge-Kutta methods [29, 21] in terms of a matrix \( A \in \mathbb{R}^{s \times s} \) and completion vector \( b \in \mathbb{R}^s \), with \( c = A \mathbf{1} \) determining the abscissa (which we will see are related to REXI poles and \( \mathbf{1} \) is a column vector of ones). The coefficients are selected to achieve the desired order of accuracy and stability properties as well as solution procedure, such as explicit, diagonally implicit, and fully implicit.

For fully nonlinear and non-autonomous ODEs \( \frac{\partial u}{\partial t} = f(t, u) \), a Runge-Kutta method in Butcher form requires solving a system of stage equations

\[ y_s = u_n + \Delta t \sum_{j=1}^{S} A_{sj} f(t + c_j \Delta t, y_j), \quad i = 1, \ldots, S \]  

and evaluating the completion formula \( u_{n+1} = u_n + \Delta t \sum_{j=1}^{S} b_j f(t + c_j \Delta t, y_j) \). Here, \( \Delta t \) is the time step size, and \( y = \{y_j\}_{j=1}^S \) is the vector of stage solutions. For linear autonomous equations, we can choose characteristic variables, in which case \( f(t, u) = \lambda u \), and the stage equations (4.2) reduce to \( y = u + \Delta t \lambda A y \) and

\[ u_{n+1} = \left[ 1 + \Delta t \lambda b^T (I - \Delta t \lambda A)^{-1} \right] u_n, \]

where we have identified the stability function \( R(z) \approx \exp(z) \).

4.1.1. Derivation. We now show that unified REXI is algebraically equivalent to Runge-Kutta methods with a diagonal Butcher matrix \( A \), starting with a decomposition inspired by the solution method developed independently by [3, 1]. Given an eigendecomposition \( A = E \Lambda E^{-1} \) (which exists for the collocation methods we will consider [16]), we can rewrite (4.3) as

\[ u_{n+1} = \left[ 1 + \Delta t \lambda b^T E (I - \Delta t \lambda D)^{-1} E^{-1} \right] u_n. \]
With \( W = \text{diag}(E^{-1})^{-1} \), we may transform to
\[
(4.5) \quad u_{n+1} = \left[ 1 + \Delta t \lambda b^T EW^{-1} (I - \Delta t \lambda D)^{-1} W E^{-1} 1 \right] u_n,
\]
which is a diagonal Runge-Kutta method with \( A \) replaced by \( D \) and the original completion vector \( b \) replaced by \( \tilde{b} \). Rewriting this to a REXI form leads to
\[
(4.6) \quad u_{n+1} = u_n + \Delta t \tilde{b}^T \left( - (\Delta t D)^{-1} \right) \left( I + (\Delta t \lambda D - I)^{-1} \right) 1 u_n
= \left( 1 - \tilde{b}^T D^{-1} 1 \right) u_n + \left( -\tilde{b}^T D^{-2} \right) \left( \Delta t \lambda - \frac{D^{-1}}{\text{diag}(\alpha)} \right)^{-1} 1 u_n.
\]
Finally, we can write this in the unified REXI formulation (1.3) with
\[
(4.7) \quad \gamma = 1 - \tilde{b}^T D^{-1} 1 \quad \beta^T = -\tilde{b}^T D^{-2} \quad \alpha = \text{diag} \left( D^{-1} \right).
\]
We have derived a transformation from implicit RK method with nonzero eigenvalues to REXI form with the same stability function. Given a REXI method, one can construct an equivalent diagonal RK method (with complex coefficients) via \( D = \text{diag}(\alpha)^{-1} \) and \( \tilde{b}^T = -\beta^T D^2 \). Note that a conventional Butcher table \( A, \tilde{b}^T \) is not uniquely determined by this procedure. We remark that standard techniques for analyzing Runge-Kutta methods can readily be applied to REXI methods. This includes barriers such as Theorem 4.3 of [20], which establishes that diagonal (parallel) RK methods can be no more than second order accurate for nonlinear problems.

**4.1.2. Error studies.** We choose the Gauss-Legendre and Chebyshev quadrature points for the error studies, with results given in Figure 4.1. We can observe that increasing the number of stages in the non-diagonalized version (using a dense Butcher table) always improves accuracy per stage. In contrast, B-REXI accuracy degrades when too many stages are used, becoming apparent beyond 8 stages. This effect is related to ill-conditioning that can be interpreted via the condition number of the eigenbasis \( E \) that effects diagonalization (4.5) or via the 1-norm of the completion vector \( \tilde{b} \), as shown in Figure 4.2. Note that completion vectors must sum to 1 so \( \|\tilde{b}\|_1 = 1 \) is optimal (and indeed holds for the original completion vector \( b \));
4.1.3. Relation to Crank-Nicolson. Since this will be relevant for the results section, we would like to show the relation between the B-REXI approximation with a single pole using the Gauss-Legendre quadrature using just a simple pole (centered at the interval). This will lead to terms \( \gamma = -1, \alpha = 2, \beta = -4 \) which yields the REXI approximation \( \exp(x) \approx -1 + \frac{4}{x+2} = \frac{1+\frac{x}{2}}{1-\frac{x}{2}} \) with the equation on the right hand side matching the Crank-Nicolson formulation. This REXI approximation with a single term resembles the Crank-Nicolson formulation with a midpoint rule (forward Euler on nominator and backward Euler on denominator for \( x = \Delta t \) and a half-time step size). This will explain that later numerical results with B-REXI match the Crank-Nicolson method. Using more REXI poles will result in even higher-order approximations.

4.2. T-REXI: Terry’s Rational Approximation of the Exponential Integrator. The approach which we will refer to as T-REXI was introduced in [18]. Several steps are required to gain the \( \alpha \) and \( \beta \) coefficients. Since these steps account for the computational workload and the properties, we briefly describe the derivation, including a discussion on the advantages and limitations of this method.

4.2.1. Derivation. The first step consists of an approximation of a Gaussian basis function \( \psi_c(x) \) as follows:

\[
\psi_h(x) = (4\pi)^{-\frac{1}{2}}e^{-x^2/(4h^2)} \approx Re \left( \sum_{k=-W}^{W} \frac{\omega_k}{i\frac{x}{h} + (\mu + ik)} \right) \tag{4.8}
\]

Using \( W = 11 \), hence \( L = 2W + 1 = 23 \) terms in total, is sufficient for an accurate approximation up to numerical double precision (see [18]). The advantage of this representation is an efficient representation of the Gaussian basis function in Fourier space. The proxy with the Gaussian basis function allows for computing the coefficients \( \nu_k \) for an approximation of an oscillatory function within an approximate range.
\[ x \in [-Mh; Mh] \] in Fourier space, yielding

\begin{equation}
\exp(ix) \approx \sum_{k=-M}^{M} \nu_k \psi_k(x + kh).
\end{equation}

Both steps are then combined, resulting in the approximation

\begin{equation}
\Re \left( \exp(ix) \right) \approx \sum_{n=-M-W}^{M+W} \Re \left( \beta_n \alpha_n^{-1} \right)
\end{equation}

where we only showed the \( \Re \) one. We combine \( \Re \) and \( \Im \) to the form

\begin{equation}
\exp(ix) \approx \sum_{n=-M-W}^{M+W} \beta_n (ix - \alpha_n)^{-1}
\end{equation}

eventually leading to the REXI formulation with \( \gamma = 0 \), but \( x \) related directly to the imaginary value on the complex plane. So far, we only targeted the \( \varphi_0 \) function, and we like to point out that this method can also be used to approximate other \( \varphi_i \) terms (see [18]) or directly with the REXI coefficients (see §3.2). We want to emphasize that this approximation was derived only for purely oscillatory functions and, hence, does not include approximations with non-zero real eigenvalues components.

### 4.2.2. Error studies
We investigate the errors of the T-REXI method in Figure 4.3. On the left image, we can observe that we need a minimum number of Gaussian basis functions to approximate the oscillations. The right image shows exceptionally accurate results for \( h \approx 0.8 \) in the range \( x \in [0; 10] \) and a rather large region of accuracy of about \( e(x = 128) \leq 10^{-11} \). Other figures (not included) show that increasing \( M \) leads to a linear increase of the size of the region of high accuracy (see [18]) with an optimum value of \( h \approx 0.8 \). For the remainder of this work, we will use \( h \approx 1.0 \) as a compromise between accuracy and total workload.

### 4.3. CI-REXI: Cauchy Contour Integral method
Cauchy Contour Integral (CI) methods offer an alternative way to infer the REXI coefficients (see e.g. [42, 4, 34]). We start with the general CI equation given by

\begin{equation}
g(x) = \frac{1}{2\pi i} \oint_{\Gamma} \frac{g(z)}{z-x} \, dz
\end{equation}
where \( g(x) \) is one of the analytic \( \varphi_i \) functions (2.4), \( \Gamma \) the contour enclosing the eigenvalue \( \lambda \) for ODEs and all eigenvalues on the diagonal of \( \Lambda \) for PDEs.

### 4.3.1. REXI Derivation.
Regarding the contour, we can use different approaches. In what follows, we used parametrized contours \( \Gamma = \{ \sigma(w) | w \in [0;1]\} \) with the contour function \( \sigma(w) : \mathbb{R} \rightarrow \mathbb{C} \). Using integration by substitution and the contour function, we obtain

\[
g(x) = \frac{1}{2\pi i} \int_{\gamma} \frac{g(\sigma(w))}{\sigma(w) - x} \sigma'(w) dw = \int_{\gamma} \frac{i(2\pi)^{-1}g(\sigma(w))\sigma'(w)}{x - \sigma(w)} dw.
\]

Using the trapezoidal rule, which is exponentially fast converging on periodic boundaries (see [41]) with \( N \) trapezoidal points in total, we obtain

\[
g(x) \approx \frac{1}{N} \sum_{n=1}^{N} \frac{i(2\pi)^{-1}g(\sigma(w_n))\sigma'(w_n)}{x - \sigma(w_n)} \quad \text{with} \quad w_n = \frac{n}{N}.
\]

Again, we can infer a unified REXI formulation (1.3) by setting

\[
\alpha_n = \sigma(w_n) \quad \beta_n = \frac{i g(\sigma(w_n))\sigma'(w_n)}{N^{2\pi}} \quad \gamma = 0.
\]

An ellipse contour is given by \( \sigma(w) = R_x \cos(iw2\pi) + iR_y \sin(iw2\pi) - \mu \) with \( \mu \) related to the center of the ellipse. This leads to the coefficients

\[
\alpha_n = R_x \cos(iw2\pi) + iR_y \sin(iw2\pi) - \mu \\
\beta_n = \frac{i}{N} \exp(\sigma(w)) (-R_x \sin(iw2\pi) + iR_y \cos(iw2\pi))
\]

A study of all kinds of contour shapes (rectangle, bean, polygonal shapes, etc.) is beyond the scope of this work. In the next section, we will mainly focus on the circle to show interesting characterizations and use the ellipse for numerical studies to show its superiority to another REXI method. In the following, we will refer to the special case of a circle as CI-REXI and to the ellipse case as CI-EL-REXI.

#### 4.3.2. Characterization and numerical issues.
Next, we characterize the REXI terms, referred to as the \( \beta \) characterization, with an overview in Figure 4.4.

We remind the reader that REXI approximates functions with a linear combination of rational basis functions. Depending on the placement of these functions (related to \( \alpha_n \)) and the weighting of each basis (related to \( \beta_n \)), we have three different cases:

a) **Obsolete REXI terms**: Contours \( Re(\sigma(x)) \rightarrow -\infty \) relating to areas of the contour in the distant negative real axis on the complex plane have exponentially fast decaying \( \beta \) coefficients, hence \( \lim_{\sigma \rightarrow -\infty} \beta_n = \lim_{\sigma \rightarrow -\infty} \frac{i\exp(Re(\sigma(w_n)))\sigma'(w_n)}{N^{2\pi}} = 0. \)

Once a particular \( \beta_n \) coefficient undershoots a threshold \( \epsilon_\beta \), the corresponding REXI term can be removed if \( \beta_n < \tau_\beta \) and \( \tau_\beta = \epsilon_\beta/N. \) The last equation incorporates that a higher numerical resolution results in smaller values of the \( \beta \) weights.

b) **Regular REXI terms**: This characterization refers to those REXI terms that can be incorporated in the approximation in a useful way.

c) **Cancellation-prone REXI terms**: These terms are related to the contour \( Re(\sigma(x)) \rightarrow +\infty \). Approximating the exp function in the far distance to the right of the origin leads to exponentially increasing the \( \beta \) values. An oscillatory function is also parallel to the imaginary axis, which is approximated. Both effects lead to
very large positive and negative numbers, resulting in severe cancellation errors in
this region. Consequently, this region should be avoided.

Examples of different contours are provided in Figure 4.5. Each contour targets a
particular problem. The circle can be used for the approximation of a small spectral
radius $\lambda \Delta t < 10$. Once requiring a larger approximation along the imaginary axis, the
radius cannot be enlarged without sacrificing accuracy due to cancellation errors in
$\beta_n$, see (c) above. This can be avoided by enlarging the radius and choosing the value
$\mu$, hence shifting the circle, to exclude a contour across areas with $Re(x) > 10$, which
leads to the shifted circle. Other contours are, e.g., given by the ellipse or rectangle
targeting the approximation of a spectrum on or close to the imaginary axis and
the bean contour targeting an approximation of diffusive and oscillatory problems.
Studies about these contours are beyond the scope of this work and we will focus on
the (shifted) circle and ellipse throughout the remainder of this paper.

4.3.3. Error studies. We conduct error studies using the shifted circle CI-REXI
method. The first study is based on a circle centered at the origin, with studies for
different radii. The second is for a circle which is shifted to overcome problems related
to the cancellation effects (see (c) above).

Results are given in Figure 4.6 with plots based on a fixed number of $N = 256
REXI poles. We can observe that the errors significantly increase for the centered
circle once the radius exceeds a certain threshold. In particular, errors for a larger
radius – including a larger spectrum on the imaginary axis – are outside the plotting
range. The results for using a higher number of REXI poles do not significantly
improve the results. Using a shifted and enlarged circle, we can gain improved results
that overcome cancellation errors.

So far, we only investigated the error itself but neglected the total workload.
Pruning $\beta$ with $\epsilon_\beta$ (exploiting characterization (a)), we can reduce some workload
significantly as depicted in Figure 4.7 for larger radii. For a moderate number of REXI
poles $N = 128$ (left image), we observe a pruning close to the accuracy of REXI itself,
hardly impacting the results. In contrast, larger radii already suffer from inaccuracies
of the used quadrature, with errors outside the plotting range. For a larger number of
REXI poles $N = 2048$ (right image), we observe very robust pruning, hardly affecting
the accuracy of the REXI approximation quality but leading to a significant reduction
of the workload.

5. Stability, normalization & filtering. So far, we have only studied errors
in approximating the $\varphi_0$ function with REXI methods. However, once we use REXI
methods for time integrating differential equations, additional properties such as sta-
bility and convergence are assumed to be relevant. We will investigate these properties
in this section for the ODE $\frac{du(t)}{dt} = \lambda u(t)$. 

This manuscript is for review purposes only.
5.1. **Stability.** The stability plots are generated based on the stability function $R(\lambda)$, which is defined by the execution of a single-time step $u(t + \Delta t) = R(\Delta t \lambda)u(t)$.

We will plot the amplification factor $|R(\lambda)|$ of the solution $u(t)$ over a time step $\Delta t = 1$.

**B-REXI** (left image in Figure 5.1): The stability reflects the A-stability of these methods on the entire left half plane. In particular, stability is given for the entire imaginary axis, a known property of collocation methods.

**T-REXI** (right image in Figure 5.1): We can observe that T-REXI provides excellent stability for purely imaginary values. However, we can observe instabilities on the imaginary axis once we reach the boundaries of the approximation range. This can be avoided by an additional T-REXI filter, which could be applied to obtain stability also outside the approximation range (see [18]).

**CI-REXI**: Finally, we look at the CI-REXI method based on Cauchy contour integral methods in Figure 5.2. The left image shows an unstable region along the imaginary axis. This is caused by an $\alpha$ pole directly placed on the imaginary axis. We can avoid this by choosing the support points of the trapezoidal rule differently. The right image shows a solution to this by shifting them by a half interval, effectively avoiding this instability, and CI-REXI becomes unconditionally stable for oscillatory systems. To summarize, if using the CI-REXI method, one should avoid placing poles...
near the eigenvalues of the linear operator.

A contour plot comparing CI methods with an increasing number of poles and approximation range is provided in Figure 5.3. We adopted the contour of the circle to pass through the points $\pm 10i$, $\pm 20i$, and $\pm 40i$ on the imaginary axis for increasing the number of poles while keeping the contour never exceeding 10 on the real axis.

5.2. Normalization. This section concerns particular problems for stationary or nearby modes requiring special treatment with the T-REXI method. So far, we only assessed errors for a single time step, and this section will investigate the accuracy and conservation properties of stationary modes concerning REXI methods. We will use Dahlquist’s equation (5) with $\lambda = 10^{-3}i$, which is time-integrated until $t = 100$ using different REXI methods. The particular choice of this low frequency is related to almost stationary modes of PDEs. Such modes play an important role, e.g., for geostrophic balance in atmospheric simulations, and not preserving them might lead to spurious/parasitic modes.

An investigation of the results at the absolute ODE errors at $t = 100$ is given in Figure 5.4. The left column shows REXI methods as they have been computed with the methods from before. We can observe that the CI-REXI method (top left image) has REXI coefficients preserving the stationary modes. However, the T-REXI suffers from significant defects in it. A normalization can be used to overcome this problem where stationary modes require $\sum_n \beta_n x^{-\alpha_n} = s = 1$ and we can ensure this by simply rescaling $\beta_n$ so that $\beta_n^{\text{new}} = \frac{\beta_n}{s}$.

The results for this are given in the right column, where we observe relatively small improvements for the CI-REXI method (right top image). However, for the T-REXI method, the errors significantly drop from $10^{-8}$ to about $10^{-13}$ once applying this normalization. We also do not see any accuracy degradation for very large time step sizes. Hence, this normalization can be used without impacting the accuracy of other choices of $\lambda$, and we will use it throughout the remainder of this work.

We close with two side comments: First, studies for purely stationary modes ($\lambda = 0$, not shown here) showed that the errors are increasing for smaller time step sizes, but only due to round-off errors. Overall, these results still lie within the range of numerical precision; hence, we skipped them here. Second, we skipped the B-REXI method since it is not prone to this problem for a number of REXI terms usable as solvers.
This manuscript is for review purposes only.
full exploration of all parameter combinations is obviously not possible. Hence, we focused on the ones that were most rational to us based on far more experiments than shown here. We first continue with concrete examples using a linear oscillatory ODE based on the Dahlquist equation followed by a PDE with the nonlinear shallow-water equations on the rotating sphere to gain insight into numerical properties once we apply this to more realistic test cases.

Based on the eigendecomposition, we classify linear operators as oscillatory, diffusive, or both. A purely oscillatory system [10] requires imaginary-only eigenvalues \((i\lambda \in \mathbb{R})\) whereas a diffusive behavior is based on negative real eigenvalues \((\lambda \in \mathbb{R} \text{ and } \lambda < 0)\). Since oscillatory/hyperbolic systems belong to the most challenging problems for REXI methods, we will solely focus on them.

6.1. ODE. We investigate the ODE systems again with Dahlquist’s equation (5) using \(\lambda = 1\) and the simulation results at \(t = 100\). We use \(u(0) = (1 + i)\sqrt{2}\) as an initial condition. We compare various REXI methods in Figure 6.1. The total numbers of REXI coefficients are given by \(N\).

The B-REXI method (left upper image) performs extremely well for small step sizes where only a few poles are required. For larger time step sizes of \(\Delta t \approx 10\), using 16 poles is sufficient to gain single precision accuracy.

The CI-REXI method (right top image) is tuned with a contour never exceeding...
a real value of 10 and to include the points on the imaginary axis given by \( Im(\lambda_{\text{max}}) \).

The CI-REXI method clearly outperforms the B-REXI method for medium-sized time step sizes and also allows taking very large time step sizes.

The T-REXI method (left bottom image) requires a significant number of REXI poles if only small time step sizes should be taken. This improves once larger step sizes are taken since the initial overheads of the large number of poles (due to the rational approximation of the Gaussian) have less relative impact on the number of REXI coefficients.

In addition, we also investigated the CI-EL-REXI (right bottom image) method, which is a natural choice for purely oscillatory problems. We chose the semi-major axis of the ellipse along the real axis in an empirical way and never exceeding 10 to avoid numerical issues. This method outperforms both CI-REXI and T-REXI almost everywhere regarding accuracy and number of terms required to solve it.

### 6.2. PDE example

In this final section, we will investigate different REXI methods with the shallow-water equations (SWE) on the rotating sphere. We decided not to investigate many different PDEs, but to go into depth of exponential integration for a single one which is of purely hyperbolic nature. We chose the SWE since they are frequently used to assess the quality and performance of discretizations in time and space concerning horizontal aspects of the full Euler equations solving the fluid dynamics equations related to the atmosphere. In velocity form, the nonlinear SWE are given by

\[
\frac{\partial}{\partial t} \begin{pmatrix} \Phi \\ \vec{V} \end{pmatrix} = \begin{pmatrix} -\Phi \nabla \cdot \vec{V} \\ -\nabla \Phi \end{pmatrix} + \begin{pmatrix} 0 \\ -f \vec{k} \times \vec{V} \end{pmatrix} + \begin{pmatrix} -\nabla \cdot (\Phi \vec{V}) \\ -\vec{V} \cdot \nabla \vec{V} \end{pmatrix}
\]

\( L_g: \) linear gravity \( L_c: \) linear Coriolis \( N: \) nonlinear term

with the horizontal velocity \( \vec{V} \) on the longitude/latitude field, geopotential \( \Phi = g \cdot h \) with height \( h \), average geopotential \( \Phi = g \cdot \bar{h} \) with average height \( \bar{h} \), a linearization around a state \( \bar{h} = 10^5 m \), Coriolis effect \( f = 2\Omega \sin(\phi) \) with latitude \( \phi \) and angular rate of rotation \( \Omega \). We like to emphasize that no (hyper)viscosity is used in this PDE to avoid a simplification of the problem due to diffusive effects.

We use this PDE due to its particularly interesting features: The linear gravity term \( L_g \) is the stiffest one and can be solved with exponential integrators either analytically or with REXI. We want to point out that a comparison of some methods has already been under investigation in former work [34] but solely with the CI-REXI method and the geopotential field, which has also been identified to be the best Strang-split method. Anyhow, this study also lacked comparisons with other variables, particularly other REXI methods, which will lead to new revelations, as presented in the following sections. Since including the T-REXI method would not provide any beneficial insight, since the CI-REXI method is computationally much cheaper and provides additional benefits, we skip this method in the following studies.

#### 6.2.1. Spatial discretization

We solve these equations using the SWEET software\(^1\) which utilizes spherical harmonics (SH) to solve these equations. Such a global spectral basis leads to a substantial reduction of spatial errors (besides a lack of nonlinear interactions at the limit of resolution), hence allowing us to put the focus on time integration methods. We like to refer to [32, 15] for a detailed description of the spherical harmonics. In particular, we work with the vorticity-divergence formulation

\(^1\)https://sweet.gitlabpages.inria.fr/sweet-www/

This manuscript is for review purposes only.
in spectral space to avoid spurious modes if one would convert the velocity to spectral space. The standard $\frac{2}{3}$ rule [27] is used for anti-aliasing to evaluate bi-non-linearities.

6.2.2. Time stepping solvers. We can find highly efficient solvers in spherical harmonics space for direct exponential integration (without numerical approximations), REXI, and implicit Euler time integrators.

Regarding the direct exponential integration, we can straightforwardly find a direct solution using the vorticity divergence form, see also [38]. Due to orthogonality, each mode can be separately written as

$$\begin{bmatrix} \partial_t \Phi' \\ \partial_t \delta \end{bmatrix} = \begin{bmatrix} -\nabla^2 & -\hat{\Phi} \\ -\hat{\Phi} & \delta \end{bmatrix} \begin{bmatrix} \Phi' \\ D & G \\ \Phi' \\ \delta \end{bmatrix} = \begin{bmatrix} \nabla^2 = -n(n+1) \end{bmatrix}$$

for this harmonic. Using $D = -\nabla^2$ and $G = -\hat{\Phi}$ for convenience, we find the eigenvectors $Q$ and eigenvalues $\text{diag}(\Lambda)$

$$Q = \begin{bmatrix} -\sqrt{\frac{G}{D}} + \sqrt{\frac{G}{D}} \frac{1}{2} \\ -\frac{1}{2} \sqrt{\frac{D}{G}} + \frac{1}{2} \sqrt{\frac{D}{G}} \end{bmatrix} \quad Q^{-1} = \begin{bmatrix} \frac{1}{2} \sqrt{\frac{D}{G}} \frac{1}{2} \sqrt{\frac{D}{G}} \\ -\frac{1}{2} \sqrt{\frac{D}{G}} - \frac{1}{2} \sqrt{\frac{D}{G}} \end{bmatrix} \quad \Lambda = \begin{bmatrix} -\sqrt{DG} \quad \sqrt{DG} \end{bmatrix}.$$  

We can then use $U(t + \Delta t) = Q \exp(\Delta t \Lambda)Q^{-1}$. From an algebraic perspective, this method matches the method in [17], which uses a rather cumbersome derivation using Laplace transforms, whereas our derivation is more elegant and short. This method is also used for investigating errors.

For the exponential integration of the full linear terms $L = L_g + L_c$, this would relate to the Hough modes [44] and no direct exponential solution has been derived yet. Hence, it requires evaluations of the form $(\Delta t L - \alpha)^{-1}$ with complex-valued $\alpha$. The first time this was solved for REXI using spherical harmonics was based on a method requiring transformations to grid space [32]. The present work is based on an implicit time stepper [37] of the form $(I - \Delta t L)^{-1}$ which has been transformed to solve a REXI term by simply using a complex-valued time step size. We also used it for the implicit time integration of $L$ as it has been originally suggested.

6.2.3. Benchmark. Our benchmark is based on the barotropic instability test case (see [14]). This benchmark is initialized with a geostrophically balanced initial
A GENERALIZED REXI FOR MASSIVELY PARALLEL TIME INTEGRATION

<table>
<thead>
<tr>
<th>Short notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERK($X, \sigma = N$)</td>
<td>Explicit Runge-Kutta with order $N$</td>
</tr>
<tr>
<td>IRK($X$)</td>
<td>Backward Euler using 2nd order Crank-Nicolson</td>
</tr>
<tr>
<td>SS($X,Y$)</td>
<td>2nd order Strang-splitting as explained in the text with $F_1 = X$ and $F_2 = Y$</td>
</tr>
<tr>
<td>EXP($X$)</td>
<td>Direct exponential integration on $X$</td>
</tr>
<tr>
<td>REXI($X$)</td>
<td>A particular REXI method on $X$</td>
</tr>
<tr>
<td>ETDRK($X,Y$)</td>
<td>2nd order ETDRK method with $X$ being exponentially integrated and $Y$ treated as the nonlinearity</td>
</tr>
</tbody>
</table>

| $X + Y$ | Time tendencies of terms $X$ and $Y$ are added |

Table 6.1

Overview of time integration methods. Note that they can be composed together.

6.2.4. Time integration. Regarding the particular Runge-Kutta (RK) based time integrators, we used 2nd order midpoint, 3rd order Heun, and classical 4th order RK. The reference solution to compute the errors is based on the 4th order RK with a time step size of $\Delta t = 5$.

Besides the methods already introduced, our investigation also includes the 2nd order Strang splitting (SS) method [36]. With SS, a PDE given by two terms $\frac{d}{dt}U = F_1(U) + F_2(U)$ can be integrated with 2nd order accuracy if a 2nd order accurate time integrator $R^{\Delta t}_{F_i}$ is provided for time step size $\Delta t$ by $U(t + \Delta t) = R^{\frac{1}{2}\Delta t}_{F_1} \circ R^{\Delta t}_{F_2} \circ R^{\frac{1}{2}\Delta t}_{F_1}$.

We use a function-like notation to refer to the particular time integration methods. An overview of this is given in Table 6.1 where we use $X$ and $Y$ as representatives for either term in the PDE such as $L_g$, $L_c$, and $N$ or to refer to another time integrator. In the latter case, e.g., ERK, EXP, REXI, and IRK can both be used in the Strang-Splitting SS as arguments.

6.2.5. Hardware, parallelization & batch configuration. All results have been computed on the Thin Nodes of SUPERMUC-NG. Each node is equipped with two Intel SkylakeXeon Platinum 8174, resulting in two NUMA domains. For the spatial parallelization, we use solely OpenMP on one NUMA domain, resulting in a spatial scaling of up to 24 cores. Scalability for REXI is then based on MPI first by utilizing the 2nd NUMA domain, then other compute nodes. We gratefully acknowledge the usage of the SHTNS library [30] which is based on FFTW [13]. We precomputed transformation plans and reused them for all studies to ensure the utilization of the same ones over all studies. Each batch job is set to timeout after 1 hour, which follows the idea that the simulations should be finished within a specific time frame.

6.2.6. Performance comparison for splitting $L_g$ and $L_c + N$. We start with a comparison of standard methods in Figure 6.3 which we will use as a baseline for further comparisons with REXI-based methods. Plots are given for the three prognostic variables, which we define here as the variables required as input to one time step, since results differ for all of them.

First, the higher-order 3rd- and 4th-order RK method can outperform other lower-order methods for smaller time step sizes depending on the variable under study.
is a known phenomenon for higher-order time integration methods, and we wanted to include it to also see its max. stable time step size. We are primarily interested in very large time step sizes while still having a moderately small error.

The method concerning the geopotential and the divergence variable is the Strang-split $SS(ERK(L_c+N), EXP(L_g))$, which we account for by the more accurate treatment with the exponential treatment of both variables. Since the vorticity field is not treated exponentially (time tendency for this in $L_g$ is null), there's also no benefit visible in the comparison of the vorticity field.

The ETDRK method itself – although assumed to be an excellent off-the-shelf method – does not provide the overall best results compared to the rather straightforward Strang splitting. We can observe it to be the 2nd best for the geopotential and even lower ranked for the other variables. We account for that by the way a 2nd order accurate Strang-splitting is performed. This can be interpreted as a subcycling of time steps by executing two half-time steps for one of the terms (the time step size limiting one).

Next, we will continue with REXI studies by comparing them with the best Strang-exponential and implicit methods from the previous results in Figure 6.4. Overall, we can observe a 2nd order convergence even if using only a single pole for the B-REXI methods.

Matching results for $SS(REXI, ERK) B-REXI N=1$ and $SS(IRK, ERK)$ are observed which is explained in §4.1.3: This particular B-REXI method resembles exactly the Crank-Nicolson method but uses one complex-valued pole to solve the system of equations.

Fig. 6.3. Studies without REXI methods (but using direct exponentiation) on all prognostic variables with error vs. time step size for the barotropic instability benchmark. We also include 2nd, 3rd, and 4th order Runge-Kutta based methods with gray lines.
The other B-REXI methods outperform all alternatives except for the direct exponential integration EXP. We see one particularly interesting and highly important effect: The B-REXI method does not provide any further advantages using more than \( N = 2 \) poles, the results are not further improved. A particularly important point is the comparison of the CI-REXI method with B-REXI, where absolutely no benefits are visible for \( N = 2 \) poles using CI-REXI compared to \( N = 2 \) poles using B-REXI. This clearly indicates that significant computational savings of a factor of 64 can be accomplished in this case compared to the former work.

We close this section by HPC studies in Figure 6.5. For sake of better overview, we only plotted the most promising candidates (ETDRK is worse than B-REXI methods, the explicit RK order 3 and 4 methods are better for larger wallclock times (smaller time steps), but unstable otherwise).

We start by comparing the performance of the direct exponential method EXP with the REXI method, where we would expect that the direct method is faster, which is not the case. We account for that by the direct method to be computationally more intensive (square root, exponential, etc., see §6.2.2) in order to solve for this term, whereas the B-REXI methods only require to evaluate two or 4 rational approximations. For the CI-REXI method, which requires \( N = 128 \) terms this is again different due to the higher MPI overheads resulting in a lower performance than the others.

Although the Strang-splitting method with the implicit term is computationally quite efficient to evaluate, its overall wallclock time performance is not optimal.

**6.2.7. Performance comparison for splitting into \( L \) and \( N \).** Next, we investigate the performance of REXI methods using a splitting into the linear term...
We can observe that fully explicit ERK methods actually provide excellent results due to a poorer performance than the more straightforward approach. N accuracy cannot be improved by using more than 8 parts dominate the overall errors. Hence, this does not provide any better results. The best methods are the 2nd order IRK-like methods already stagnate. With respect to the ETDRK scheme, its performance is worse compared to the best (straightforward) Strang-split methods.

For the geopotential \( \Phi \), we can observe significant improvements in terms of accuracy. In particular, we can take very large time step sizes and still observe a convergence, whereas the 2nd order IRK-like methods already stagnate. With respect to the ETDRK scheme, its performance is worse compared to the best (straightforward) Strang-split methods.

For the vorticity \( \eta \) we can observe that the ETDRK method does not lead to any improvement. The best methods are the Strang-split IRK-based ones and some REXI-based methods. Hence, we do not see any improvement in the accuracy of the vorticity field by using exponential integration methods. This is kind of surprising at first glimpse since we expected a better treatment of the vorticity due to the exponential integration of the Coriolis effect. However, the errors in the nonlinear parts dominate the overall errors. Hence, this does not provide any better results.

The divergence \( \delta \) study shows REXI methods to be the best ones. Again, the accuracy cannot be improved by using more than \( N = 2 \) poles. Everything beyond that would be an additional computational burden. The ETDRK methods again show a poorer performance than the more straightforward approach.

Finally, we investigate the wallclock time vs. errors with results given in Figure 6.7. We can observe that fully explicit ERK methods actually provide excellent results due to

---

**Fig. 6.5.** Studies including REXI methods with wallclock time vs. time step size for the barotropic instability benchmark.
to their computationally efficient way. In particular, the classical 4th-order accurate ERK method provides excellent results across all prognostic variables.

A closer look at the geopotential $\Phi'$ errors shows that the B-REXI-based methods with $N=2$ poles are to be preferred compared to all other methods. Again, the ETDRK method shows no real benefits.

Investigating the vorticity $\gamma$ leads to a different interpretation: Now, the implicit Strang-split method provides the best results which can be easily explained by the situation that the exponential treatment of the $L_t$ term did not lead to any beneficial results already in the error vs. time step size plots and additional computational time is required here. Finally, ETDRK are literally the worst in here, not paying off at all.

The errors on the divergence $\delta$ show similar results compared to the geopotential, which is why we skip a detailed discussion here.

### 6.2.8. Summary of PDE results.

The CI-REXI method with $N = 128$ poles is not beneficial at all compared to B-REXI with $N = 2$ poles. Using only $N = 2$ poles with the B-REXI method already provides the best results, and no improvement can be gained by adding more poles. This is actually quite surprising, with expectations of exponential integration methods to always provide significantly better results. However, using such a higher-order approximation seems to provide sufficient accuracy so that the errors from the splitting approach dominate the overall errors.

We would like to emphasize that all the statements are specific to the SWE on the rotating sphere PDE and should not be generalized.
7. Summary and Conclusions. Exponential integration methods are considered to be a way to integrate with high efficiency. As part of that, \( \varphi \) functions need to be solved, which turn out to be computationally rather challenging.

This paper investigated different ways to approximate \( \varphi \) functions with rational approximations of exponential integration (REXI). The coefficients of REXI methods can be derived in many ways and we introduced a generalized REXI approach, finally allowing to express many different methods in this way. We showed this for the Butcher/Bickard-based REXI, Cauchy Contour integration based REXI and T(erry)-REXI method. All methods have been introduced in a way making its capabilities and limitations easily graspable. With respect to physical properties, the T-REXI method requires special treatment for (quasi-)stationary modes and became obsolete with CI-REXI. In addition, we derived an elegant way to compute higher-order \( \varphi \) functions based on REXI coefficients for lower-order \( \varphi \).

An in-depth investigation of the approximation quality of each REXI method has been conducted including an explanation of numerical issues for all of the methods. Next, we put it into the context of time integration methods. We first used linear ODEs where we studies and discussed properties of stability, convergence and also the filtering capabilities. Second, we performed in-depth studies using the nonlinear shallow-water equations on the rotating sphere. Surprisingly, the best REXI method turned out B-REXI with only \( N = 2 \) poles, leading to a significant reduction of computational effort compared to former REXI methods in this context using \( N = 128 \) poles. Consequently, regarding demands on computational resources, B-REXI showed a reduction of a factor of 64 compared to previous work. This also means that a
higher-order implicit Runge-Kutta method is competitive to traditional exponential integration methods for this PDE.

Acknowledgements. Both authors like to thank Pedro S. Peixoto for pointing out the potential relation of exponential integration methods to Laplace transforms and Peter Lynch’s work in this context. Martin Schreiber is grateful to NCAR for providing financial support and a very inspiring office space with a splendid view to the flatirons, which strongly supported this work. Both authors thank Matthew Normile for preliminary work as well as Finn Capelle and Raphael Schilling who indirectly contributed to this work with the REXInsight software.

The authors gratefully acknowledge the Gauss Centre for SC e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputer SUPERMUC-NG at Leibniz Supercomputing Centre (www.lrz.de).

REFERENCES
