

# Parameter studies for REXI

Martin Schreiber <M.Schreiber@exeter.ac.uk>  
et al.

September 5, 2015

This documents is on results for the REXI as described in [2] [3] and is based on the linear formulation only. Several benchmarks are executed to evaluate the dependency of the REXI parameters on scenarios representative for geophysical flows.

## 1 Benchmark scenario description

### 1.1 Equations

We analyze the behavior of REXI based on parameter studies of the linear part

$$L(U) := \begin{pmatrix} 0 & -\eta_0 \partial_x & -\eta_0 \partial_y \\ -g \partial_x & 0 & f \\ -g \partial_y & -f & 0 \end{pmatrix} U$$

of the SWE (see [1], note that the sign is inverted) with

$$U_t = \epsilon L(U)$$

with

$$U := (h, u, v)^T.$$

If not otherwise stated, we set  $\epsilon = g = \eta_0 = f = 1$  and  $\Omega = [0; 1]^2$ .

### 1.2 Abbreviations

We give a brief overview of the used abbreviations for the parameter studies:

Symbol	Description
$h$	REXI parameter specifying the sampling accuracy
$M$	REXI parameter related to the number of poles
$NxN$	Resolution of simulation domain
$\tau$	Time step size for REXI
$dt$	Time step size for REXI (identical to $\tau$ in this work)
$DT$	Overall simulation time
$nT$	Number of time steps

### 1.3 Initial conditions

We consider the following initial conditions

#### 1.3.1 Scenario “Gaussian” (Default)

If not otherwise stated, we use the Gaussian function

$$\begin{aligned} \eta(x, y) &:= \eta_0 + e^{-50(x^2+y^2)} \\ u(x, y) &:= 0 \\ v(x, y) &:= 0 \end{aligned}$$

as initial conditions and use zero values for initial velocity conditions.

### 1.3.2 Scenario “Waves”

We also evaluate the following initial parameter from [2]:

$$\begin{aligned} h(x, y) &:= \sin(6\pi x)\cos(4\pi y) - \frac{1}{5}\cos(4\pi x)\sin(2\pi y) + \eta_0 \\ u(x, y) &:= \cos(6\pi x)\cos(4\pi y) - 4\sin(6\pi x)\sin(4\pi y) \\ v(x, y) &:= \cos(6\pi x)\cos(6\pi y) \end{aligned}$$

## 1.4 Discretization in time

A Runge-Kutta time stepping method of 4-th order is used if appropriate. E.g. RK4 is not used for REXI, but for the other methods discussed in the next section.

## 1.5 Solvers and their discretization in space

The tests are based on different implementations:

### 1.5.1 Used derivatives

Furthermore, the derivatives can be based on

- (FDderiv) Finite difference methods (e.g. a  $[-1, 0, 1]$  stencil) or
- (SPderiv) computing the derivative via the spectral basis functions  $(\frac{d}{dx}e^{ixj2\pi})$ .

### 1.5.2 Realization of operator

Derivatives can be expressed either in

- (Cop) Cartesian space via a stencil computation (expensive in Cartesian space) or
- (Sop) Spectral (Fourier) space via a spectral convolution (element-wise multiplication and hence very cheap in spectral space).

### 1.5.3 Grid alignment

- (A-grid): All conserved quantities are placed at the cell center
- (C-grid): The potential is placed at the cell center and the velocity components at the cell edges

### 1.5.4 REXI

There is only one REXI implementation, but with different parameters which are discussed below.

## 1.6 REXI

The REXI approach allows testing different parameters. In this work, we analyze the accuracy of REXI, based on the parameters  $h$  and  $M$ , see [3] for their description. We denote the result  $U'(\tau)$  of the REXI approach by

$$U'(\tau) := \text{REXI}(U(0), \tau, h, M)$$

with  $U$  the solution at  $t = 0$ , the parameter  $\tau$  the size of the large time step,  $h$  the sampling interval size and  $M$  the number of poles.

## 1.7 Error norm

We use the RMS error norm

$$\text{rms}(f) := \sqrt{\frac{\sum_{\vec{x}} (f_{\vec{x}} - \tilde{f}_{\vec{x}})^2}{\prod_i \vec{N}_i}}$$

based on a discrete computed solution  $f_{\vec{x}}$  for the points given by  $\vec{x}_i \in \{0, 1, \dots, \vec{N}_i\}$ , the reference solution given by  $\tilde{f}(\vec{x})$  and with each component in  $\vec{N}$  giving the resolution.

## 2 Analytical benchmark results

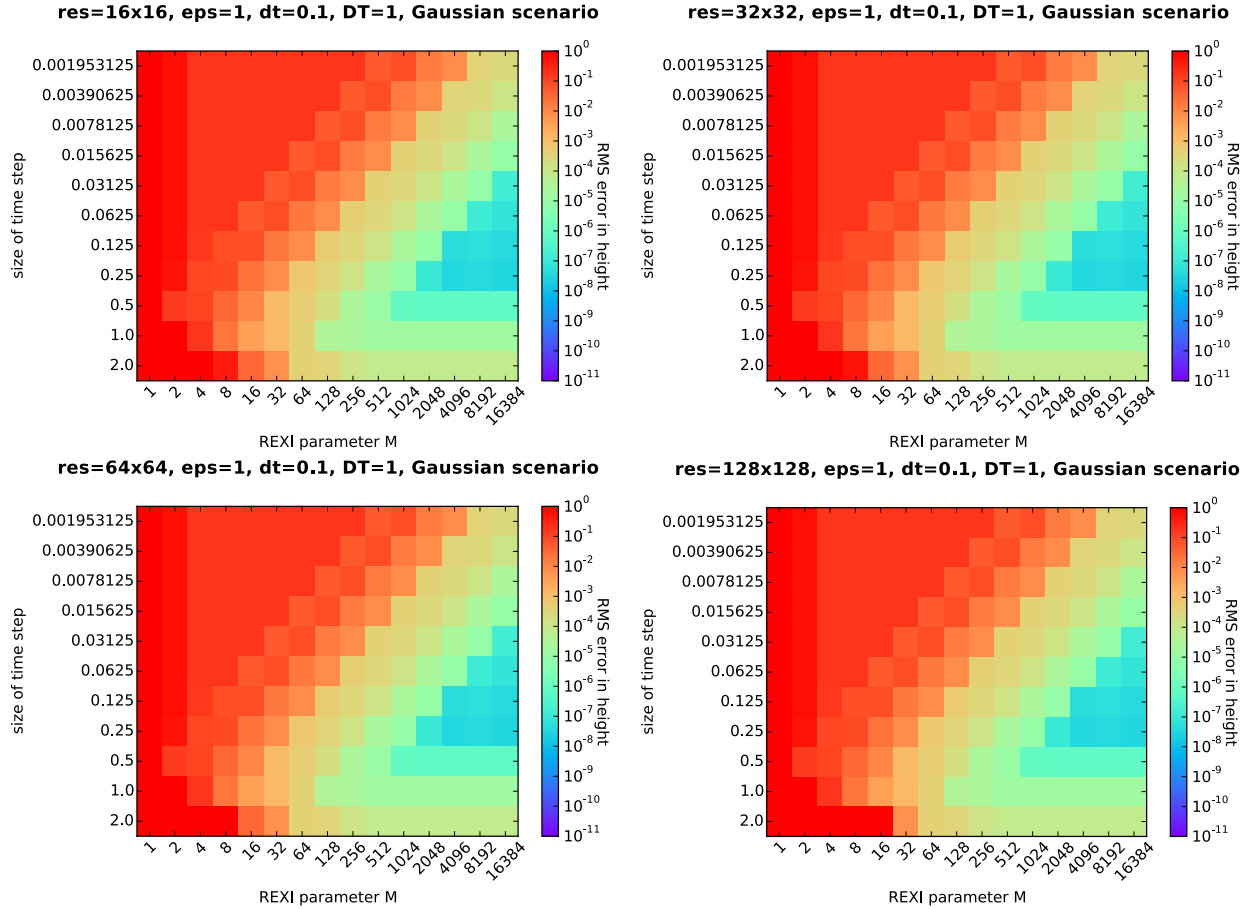
### 2.1 Study for REXI parameters $h$ and $M$

(benchmark folder: 2015\_08\_30\_search\_h\_M) [TODO: reproduce these results 2015\_09\_04\_search\_h\_M]

We first search for an optimal choice of  $h$  and  $M$  values.  $h$  is a value which only relates to the accuracy of the REXI approximation and  $M$  relates to the accuracy, as well as to the computational workload which has to be invested for the REXI approach. We execute the simulation for an overall simulation time of  $DT := 1$  second and use a time step size for REXI of  $dt := 0.1$ . Hence, 10 REXI time steps are executed. The results are given based on the computation of the RMS error on the height.

#### 2.1.1 Different resolutions

We first analyze a parameter study for different resolutions



We can observe a (blue-colored) triangle with high-accurate solutions of  $O(10^{-6})$ . This optimal triangle area for a given numerical accuracy  $e$  can be described by

$$A_{(M,h)} := \{(M, h) | Mh > C(e), h < 0.5\}$$

with  $C := 1024$  for an accuracy  $\approx 10^{-7}$ .

The diagonal edge at  $Mh = \text{const.}$  is explained by  $h$  specifying the sampling interval and  $M$  the sampling nodes. If a smaller value of  $h$  results in an oversampling of the frequency, more poles  $M$  have to be used to accurately cover the approximation of the solution. Since we should try to minimize  $M$  as far as possible, we can determine a particular best-investment for  $M$  at the left lower corner of the triangle area by choosing  $h \in [0.125, 0.25]$ .

We define such an optimal choice of  $h$  and  $M$  as the sweetspot for a given numeric solution  $U'(t + \tau, h, M) := \text{REXI}(U(t), \tau, h, M)$  and a given numerical error threshold  $e$ . We then search for parameters  $(M, h)$  which keep the error below a certain error threshold  $e$  and which minimize the number of poles directly related to  $M$ . This sweetspot is given by

$$S(h, M) := \min_M \min_h (\text{rms}(U'(t + \tau, h, M) - U(t\tau)) < e)$$

for the optimal choice of the parameters  $(M, h)$ .

### 2.1.2 Resolution $16^2$ , $32^2$ and $64^2$

Next, we test if both parameter change for different resolutions  $16^2$  and  $64^2$ :

We can observe, that  $h$  can be fixed and we use  $h := 0.2$  if not otherwise stated.

For these benchmarks, a change in resolution did not significantly change the results for selected parameters  $h$  and  $M$  and continue investigating this possible independence to the resolution further.

### 2.1.3 Summary

REXI parameter	Observation
$h$	Optimum given at $h := 0.2$
$M$	Independent to the resolution

Given a desired numerical accuracy  $e$  and with  $h := 0.2$  assumed to be constant, the optimal parameters  $M$  is given by

$$M := \frac{C(e)}{h}. \quad (1)$$

## 2.2 Testing for different $\epsilon$

(benchmark folder: 2015\_08\_30\_search\_epsilon, 2015\_09\_03\_search\_epsilon, 2015\_09\_03\_search\_epsilon\_b)

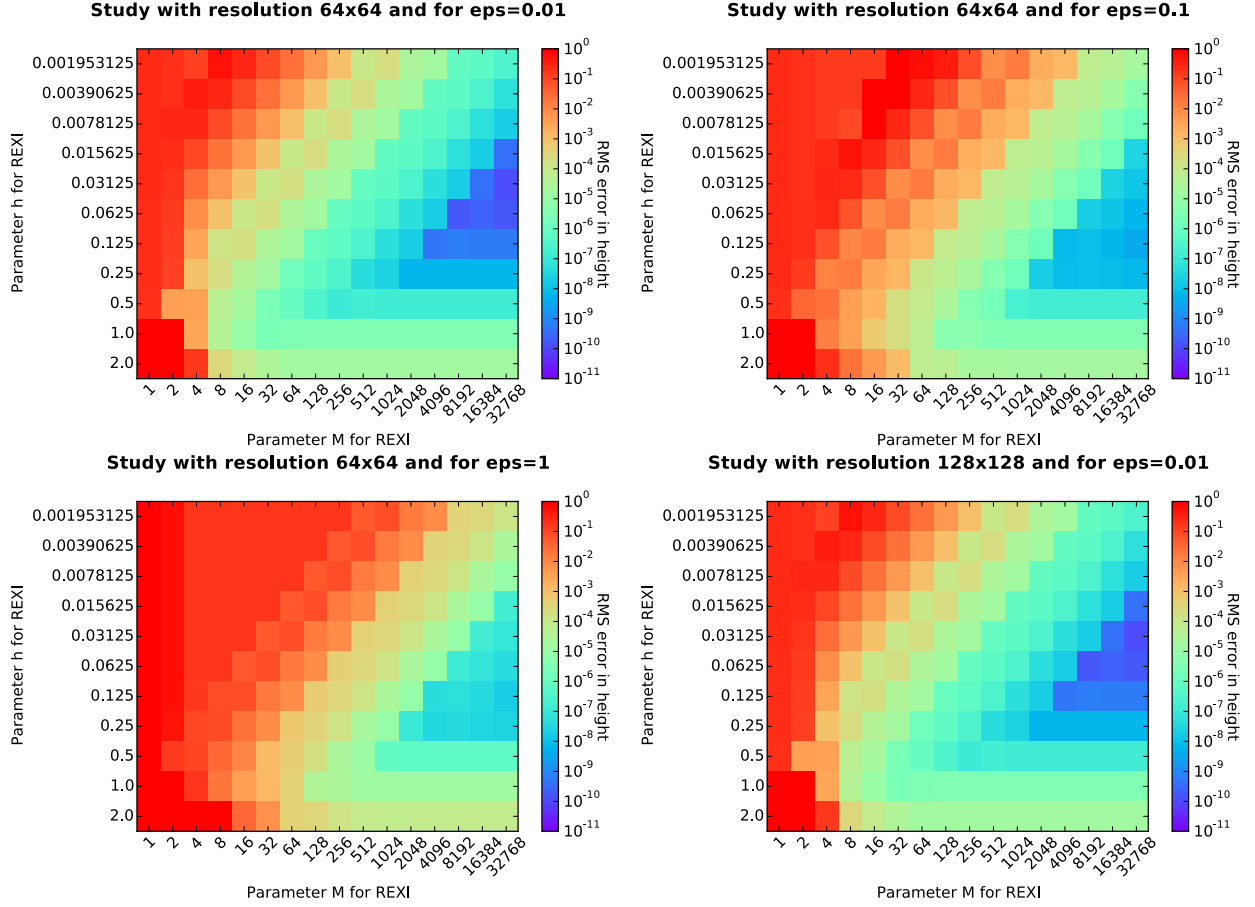
The REXI approach should hold for all the possible regimes of atmospheric constellations. Therefore, we generalize our formulation with  $\epsilon$

$$U_t = \epsilon L(U)$$

which we realize by setting  $g = \eta_0 = f = \epsilon$ .

### 2.2.1 Studies for $\epsilon := \{1, 0.1, 0.01\}$

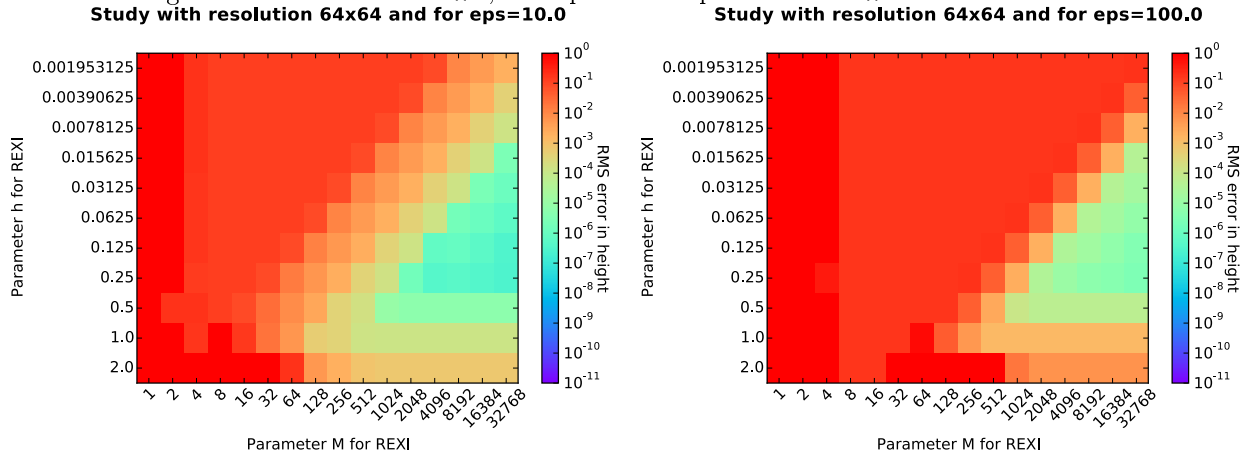
We fix our test resolution to  $64^2$  for this evaluation and show plots for  $\epsilon := \{1, 0.1, 0.01\}$ :



For a decreasing  $\epsilon$ , we can see that the previously clear optimal parameters for  $(h, M)$  are now getting more diffusive. To see if the parameters  $(h, M)$  depend on the resolution, we added a plot for benchmarks with the resolution  $128^2$  and  $\epsilon := 0.01$ , see above. We can see, that the parameters of the REXI approach are again independent to the resolution for this scenario.

### 2.2.2 Studies for $\epsilon := \{10, 100\}$

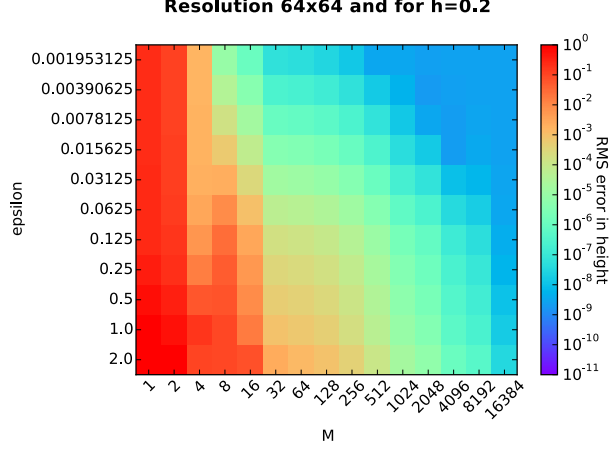
Since the borders got more diffusive for  $\epsilon \ll 1$ , we expect a sharp border for  $\epsilon \gg 1$  which we evaluate next:



However, we don't see an increase in diffusiveness. But we can observe, that the solution area  $A_{(M,h)}$  is shifted to the right side. Hence, for larger  $\epsilon$  values (this increases the stiffness of the problem), we also have to increase  $M$ . We observe again, that the solution area  $A_{(M,h)}$  still seems to be independent to the choice of  $h$ .

### 2.2.3 Dependency of $M$ to $\epsilon$

To determine the dependency of  $M$  to  $\epsilon$ , we finally executed a parameter study by setting  $h := 0.2$  and with variables  $\epsilon$  and  $M$ . Here, we keep the simulation time limited to 1 second and use a REXI timestep size of  $dt := 0.1$ .



We can see a linear dependency of  $M$  to  $\epsilon$  in the range of  $\epsilon \in [0.01; 1]$ .

### 2.2.4 Summary

Regarding the results from Section 2.2.3, the REXI parameter  $M$  is indirectly proportional to  $\epsilon$ . In combination with Eq. (1), we can extend the formula for automatically determining  $M$  to

$$M := \frac{C(\epsilon)}{h\epsilon}. \quad (2)$$

## 2.3 Study for REXI parameter $M$ and resolution $N \times N$

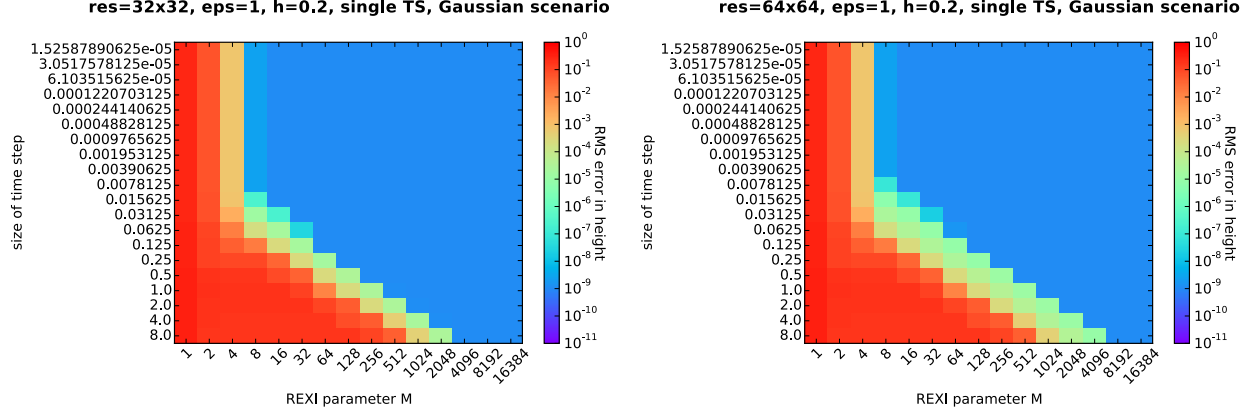
We start by evaluating the dependency of the REXI parameter  $M$  to the resolution  $N \times N$  by a single time step. We again set  $h := 0.2$  for these benchmarks and evaluate the dependency with the initial conditions described in Section (1.3).

### 2.3.1 Single time step

(benchmark: 2015\_09\_04\_increasing\_dt\_single\_timestep\_s1, 2015\_09\_04\_increasing\_dt\_single\_timestep\_s5)

The first tests are performed with a single time step and varying time step size.

We first have a look on the results for the Gaussian initial conditions:



We can observe, that there is indeed a dependency on the used resolution. However, this can be induced by using zero-valued velocity components as initial conditions. Therefore, we further evaluate this with initial conditions with non-zero velocities:

**res=6**

1.52587890625e  
3.0517578125e  
6.103515625e  
0.00012207031  
0.0002441406  
0.000488281  
0.00097656  
0.0019531  
0.003906  
0.00781  
0.0156  
0.031  
0.06  
0.1  
0

size of time step

Here, we can observe that the accuracy of the solution again is independent to the resolution. An increasing time step size  $dt$  requires a linear increase of  $M$ .

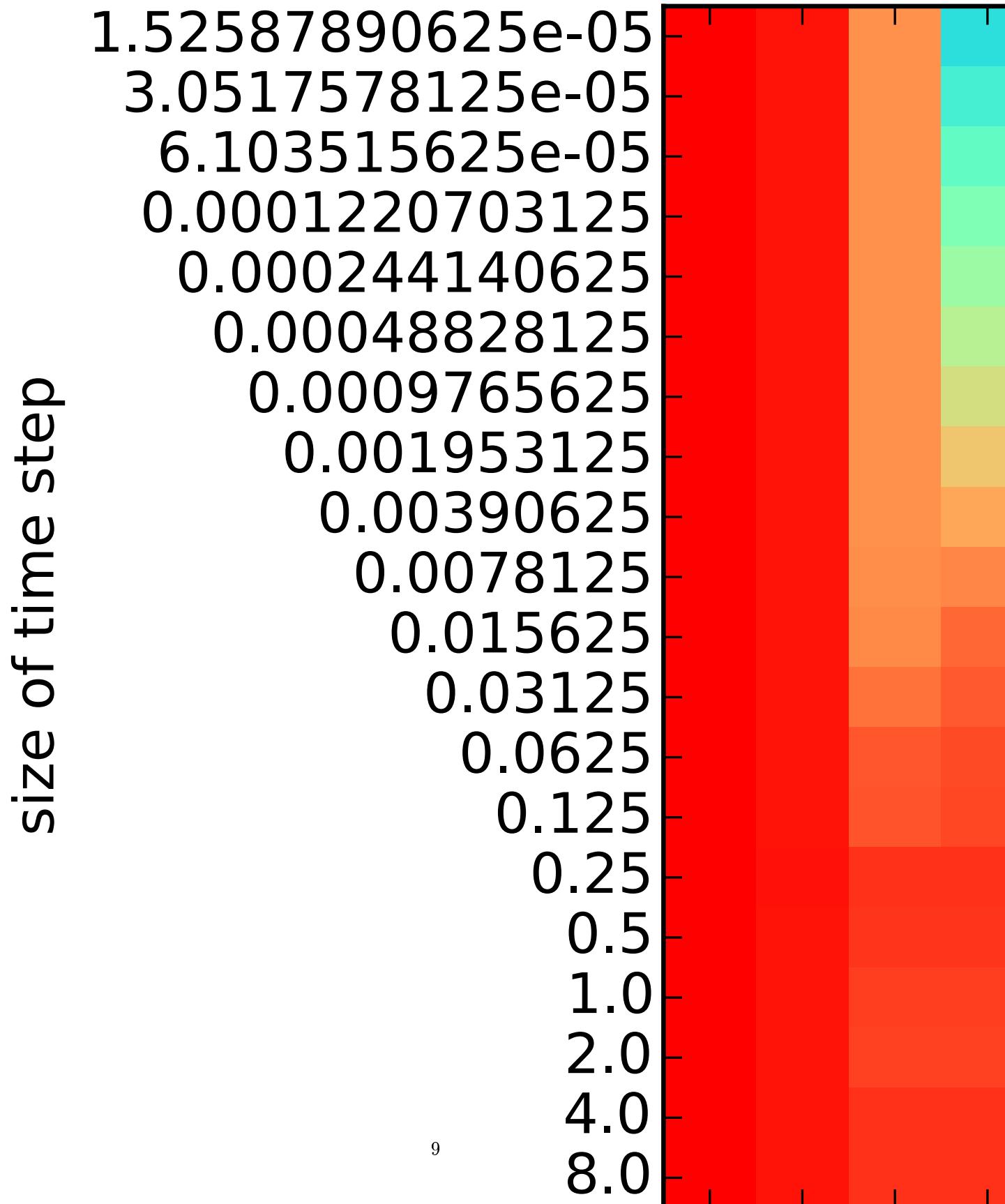
### 2.3.2 Multiple time steps

(benchmark: 2015\_09\_04\_increasing\_dt\_10\_timesteps\_s1, 2015\_09\_04\_increasing\_dt\_10\_timesteps\_s5)

Next, we analyze the accuracy over multiple time steps with the Gaussian and Waves as initial conditions:



**res=32x32, eps=**

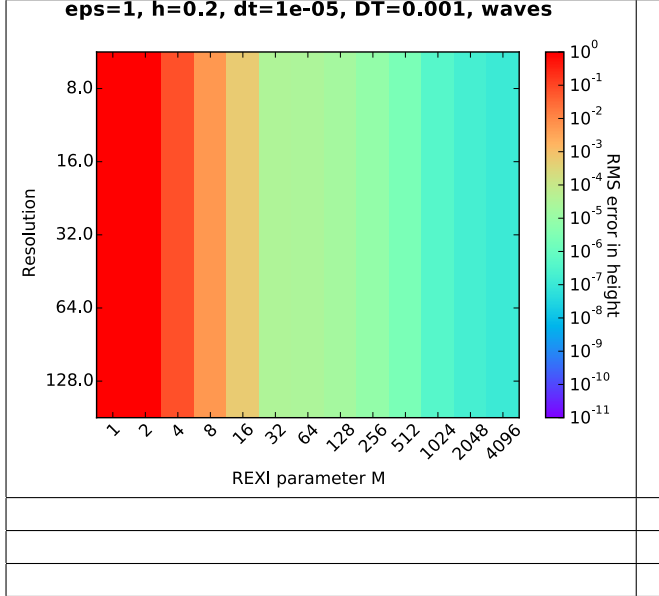


We can again observe, that the results show an independence to the resolution. Again, an increasing time step size  $dt$  results in a linear increase of  $M$ .

### 2.3.3 Different time step sizes over varying resolution $N$ and $M$

(benchmark: 2015\_09\_04\_study\_N\_M\_small\_ts)

The independence to the resolution can be caused by the large time step size dominating the error. To check that the error in time is not dominating (which would explain the independency of  $M$  to the resolution), we further analyze the error with very small time step sizes  $dt$ , varying  $M$  and varying resolution  $N$ :



For these small time step sizes, we can observe, that the change in resolution still does not result in any change regarding the parameter  $M$  depending on the resolution  $N$ .

### 2.3.4 Conclusion

An increasing time step size  $dt$  requires a linear increase of  $M$ . Therefore, we extend Equation (2) with  $dt$ .

$$M := \frac{dt C(e)}{h\epsilon}. \quad (3)$$

## 3 Performance benchmarks

Here, we compare the REXI approach with several other possible discretizations, see Sec. 1.5.

### 3.1 Performance studies for time stepping methods

We executed performance studies on a 40-core Intel Westmere system (4 sockets) to compare the overall computation time of several discretization methods with the REXI approach in both accuracy and time to solution. We tested two different parallelization concepts for the REXI method:

- REXI standard: The first one is the default one which parallelizes over the different solvers
- REXI PARSUM: The second one uses a parallelization over the sum and computes each term in the sum using only a single core.

[TODO]

### 3.2 Evaluation via a performance model

With a standard time stepping method, we denote the costs for computing a single time step by  $C_N \left( \prod_i \vec{N}_i \right)$ . Furthermore, we can assume that  $T := O\left(\frac{t \cdot \vec{v}_{max}}{N}\right)$  time steps are required for computing the solution with  $\vec{v}_{max} := \max_i(\vec{v})$  denoting the maximum velocity.

We assume that  $M$  is directly proportional to  $\vec{v}_{max}$  being related to  $\epsilon^{-1}$ .

For a given  $M$ -value, and by using the symmetry of the  $\alpha_i$  poles in REXI, this requires solving  $L + M + 1$  times the system  $(L - \alpha)^{-1}$ .

## 4 Summary

We summarize

## 5 Discussion of possible alternative approaches

### 5.1 Eigenvector decomposition in spectral space

Since the analytical solution is directly given by an Eigenvalue decomposition in spectral space, this could be an appropriate alternative to using the REXI approach. This is true for the  $f$ -plane. Let's analyze this property with the  $\beta$ -plane. Here, the  $f$  value is not constant over the plane, but varies depending on the  $y$ -coordinate:

$$f(y) := f_0 + \beta y$$

We can now have a look at the spectral formulation of the  $L(U)$  operator

$$-L(U) := \begin{pmatrix} 0 & \eta_0 \partial_x & \eta_0 \partial_y \\ g \partial_x & 0 & -f(y) \\ g \partial_y & f(y) & 0 \end{pmatrix} U$$

with  $U(\vec{x})$  given by its spectral superposition

$$U(\vec{x}) := \int_{\vec{k}} \tilde{U}(\vec{k}) e^{i\vec{k}\vec{x}}$$

and  $f(\vec{x})$  given by

$$f(\vec{x}) := \int_{\vec{k}} \tilde{f}(\vec{k}) e^{i\vec{k}\vec{x}}.$$

We were able to compute the solution to  $L(U)$  directly in spectral space with the EV decomposition, since we were able to factor out the spectral basis functions. An example is given here for the term  $\eta_0 \partial_x$ :

$$(\eta_0 \partial_x) \int_{\vec{k}} \tilde{U}(\vec{k}) e^{i\vec{k}\vec{x}} = \int_{\vec{k}} \tilde{U}(\vec{k}) \eta_0 \partial_x e^{i\vec{k}\vec{x}} = \int_{\vec{k}} \tilde{U}(\vec{k}) \eta_0 i\vec{k} e^{i\vec{k}\vec{x}} = \int_{\vec{k}} \tilde{U}(\vec{k}) \left( \eta_0 i\vec{k} e^{i\vec{k}\vec{x}} \right)$$

In combination with a Ritz-Galerkin approach and using the orthogonality of the basis functions, this allows a wave-wise Eigenvalue decomposition. However for the  $f(x)$  term, we are not able to factor out the basis function and apply the previously mentioned method

$$(f(y)) \int_{\vec{k}} \tilde{U}(\vec{k}) e^{i\vec{k}\vec{x}} = \int_{\vec{k}} f(\vec{k}) e^{i\vec{k}\vec{x}} \int_{\vec{k}} \tilde{U}(\vec{k}) e^{i\vec{k}\vec{x}}$$

which results in a non-linearity. Hence, we are not able to factor out the spectral basis functions and the approach is not applicable to the  $\beta$ -plane anymore.

## References

- [1] Formulations of the shallow-water equations, M. Schreiber, P. Peixoto et al.
- [2] High-order time-parallel approximation of evolution operators, T. Haut et al.
- [3] Understanding the Rational Approximation of the Exponential Integrator (REXI), Schreiber M., Peixoto P.