

University of Groningen

## Numerical methods for studying transition probabilities in stochastic ocean-climate models

Baars, Sven

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2019

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Baars, S. (2019). *Numerical methods for studying transition probabilities in stochastic ocean-climate models*. Rijksuniversiteit Groningen.

### Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

### Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# TRANSITIONS IN THE MERIDIONAL OVERTURNING CIRCULATION

For high-dimensional problems, the computation of transition probabilities of rare events is not feasible with standard Monte Carlo methods as mentioned in the previous chapter. Therefore specialized methods such as Genealogical Particle Analysis (GPA) (Moral and Garnier, 2005; Moral, 2013; Wouters and Bouchet, 2016), Adaptive Multilevel Splitting (AMS) (Cérou and Guyader, 2007; Rolland and Simonnet, 2015; Rolland et al., 2015), and based on that Trajectory-Adaptive Multilevel Sampling (TAMS) (Lestang et al., 2018) exist. The idea behind these methods is similar: they try to push samples in the direction of the transition to make sure transitions actually occur, also for smaller sample sizes, while still being able to compute the transition probability for the original problem.

Our goal is to compute transition probabilities in the Meridional Overturning Circulation (MOC). Rare transitions in the two-dimensional barotropic quasi-geostrophic equations have been studied in Bouchet et al. (2014); Laurie and Bouchet (2015), but there only the most likely transition path was determined by means of the minimum action method (E et al., 2004; Vanden-Eijnden and Heymann, 2008; Zhou et al., 2008; Grafke et al., 2017). To our knowledge, actual transition probabilities have never been computed for high-dimensional problems on a scale that we present here.

Methods for finding transition probabilities have two main culprits. The first is that they always require some form of time integration, which is very expensive for high-dimensional systems. This is especially the case when an implicit time-stepping method is used, since then in every time step a linear system has to be solved. The other is that for methods based on AMS, a large number of states at different time steps have to be stored in memory.

We therefore propose a projected time-stepping method in Section 6.1, for

which in every time step, only a small projected linear system has to be solved, and for which only the projected states have to be stored, which are much smaller than the original states. Based on the results that were obtained in the previous chapter, we decided to apply this method to TAMS, but it may be applied to any method for computing transition probabilities. In Section 6.2 we then discuss the problem setting and in Section 6.3 we discuss the results. The results and the method are summarized and discussed in Section 6.4.

## 6.1 Projected time-stepping in TAMS

In AMS and TAMS, but also in GPA, one is free to choose any stochastic time stepping method. In case the stochastic theta method (2.8) is used with  $\theta \neq 0$ , multiple linear systems have to be solved in every time step. In that case, solving these linear systems is usually the computationally most expensive part of TAMS, even if the linear system is sparse, and an iterative solver with a good preconditioner is used.

At a time step  $i$  of the stochastic theta method, take

$$\hat{F}(\mathbf{x}) = M\mathbf{x}_i - M\mathbf{x} + (1 - \theta)\Delta t F(\mathbf{x}_i) + \theta\Delta t F(\mathbf{x}) + g(\mathbf{x}_i) \Delta \mathbf{W}_i$$

with Jacobian

$$\hat{J}(\mathbf{x}) = \theta\Delta t J(\mathbf{x}) - M.$$

A Newton solve for one implicit time step would normally look like this

- 1:  $\mathbf{y}^{(0)} = \mathbf{x}_i$
- 2: **for**  $j = 1, 2, \dots$  **until convergence do**
- 3:     **solve**  $\hat{J}(\mathbf{y}^{(j-1)})\Delta\mathbf{y}^{(j)} = -\hat{F}(\mathbf{y}^{(j-1)})$
- 4:      $\mathbf{y}^{(j)} = \mathbf{y}^{(j-1)} + \Delta\mathbf{y}^{(j)}$
- 5:  $\mathbf{x}_{i+1} = \mathbf{y}^{(j)}$

A small potential optimization that can be made is computing  $\hat{J}(\mathbf{x}_i)$  beforehand, and using this instead of  $\hat{J}(\mathbf{y}^{(j-1)})$ . This will make it such that Newton does not converge quadratically, but saves the time of computing the Jacobian matrix multiple times.

Instead of solving with this large sparse matrix  $\hat{J}(\mathbf{y}^{(j-1)})$  in every Newton iteration, we instead propose to solve with a smaller matrix obtained from a Galerkin type projection  $V^T \hat{J}(\mathbf{x}_i) V$ , where  $V$  consists of an orthogonal basis of the most important directions that are used by the stochastic perturbation.

These vectors, at least around a steady state, can be obtained from the eigenvectors belonging to the largest eigenvalues of the covariance matrix, which can in turn be obtained from the basis vectors of the low-rank solution of a generalized Lyapunov equation as described in Chapter 4.

Say we have an orthonormal basis  $V$  spanned by  $\bar{x}_A$ , the basis of the low-rank Lyapunov solution  $V_A$  at  $\bar{x}_A$ ,  $\bar{x}_B$ , the basis of the low-rank Lyapunov solution  $V_B$  at  $\bar{x}_B$ , and the noise matrices  $B_A$  and  $B_B$  at  $\bar{x}_A$  and  $\bar{x}_B$ . We can then replace the Newton iteration with

```

1:  $\mathbf{y}^{(0)} = V^T \mathbf{x}_i$ 
2:  $A = V^T \hat{J}(\mathbf{x}_i) V$ 
3: for  $j = 1, 2, \dots$  until convergence do
4:   solve  $A \Delta \mathbf{y}^{(j)} = -V^T \hat{F}(V \mathbf{y}^{(j-1)})$ 
5:    $\mathbf{y}^{(j)} = \mathbf{y}^{(j-1)} + \Delta \mathbf{y}^{(j)}$ 
6:  $\mathbf{x}_{i+1} = V \mathbf{y}^{(j)}$ 

```

If we apply this in TAMS, we see that between two consecutive time steps, we always apply  $V^T V \mathbf{y}^{(j)}$ . Now since  $V$  is orthonormal, this means that we might as well apply TAMS itself to  $\mathbf{y}^{(j)}$  directly. Doing this actually solves the largest problem we observe when applying TAMS to high-dimensional systems, which is the storage of the trajectories, since we now only have to store the trajectories restricted to the space  $V$ . Say we have a system of size 10000 and a space  $V$  consisting of 100 vectors, which is a realistic scenario, then this reduces the required storage by a factor 100.

## Problem setting 6.2

In this section we discuss the problem setting of the ocean model as described in Section 2.5.2. The results with this model are shown in Section 6.3.

### Bifurcation diagram 6.2.1

For the deterministic model, we take the surface forcing slightly different from the forcing that was described in Section 2.5.3 and more similar to what was used in Van der Mheen et al. (2013) as

$$\bar{T}(\theta) = 10.0 \cos(\pi\theta/\theta_N), \quad (6.1a)$$

$$F_s(\theta) = \bar{F}_s(\theta) = \mu F_0 \frac{\cos(\pi\theta/\theta_n)}{\cos(\theta)} + \beta F_0 F_p(\theta), \quad (6.1b)$$

where  $\mu = 0.3$  is the strength of the mean freshwater forcing,  $F_0 = 0.1 \text{ m yr}^{-1}$  is a reference freshwater flux,  $\beta$  is the strength of an anomalous freshwater flux  $F_p$ , which is 1 in the area  $[45^\circ N, 60^\circ N]$ , and 0 elsewhere.

The equations are discretized on a latitude-depth equidistant  $n_x \times n_y \times n_z$  grid using a second-order conservative central difference scheme. An integral condition expressing the overall conservation of salt is also imposed, as the salinity equation is only determined up to an additive constant. The total number of degrees of freedom is  $n = 6n_x n_y n_z$ , as there are six unknowns per point. The standard spatial resolution used is  $n_x = 4, n_y = 32, n_z = 16$  and the solutions are uniform in the zonal direction, with the zonal velocity  $u = 0$ .

The bifurcation diagram of the deterministic model for parameters as in Table 2.1 is shown in Figure 6.1a. On the  $y$ -axis, the sum of the maximum ( $\Psi^+$ ) and minimum ( $\Psi^-$ ) values of the meridional streamfunction  $\Psi$  is plotted, where  $\Psi$  is defined through

$$\frac{\partial \Psi}{\partial z} = v \cos \theta, \quad -\frac{1}{r_0 \cos \theta} \frac{\partial \Psi}{\partial \theta} = w. \quad (6.2)$$

For the calculation of the transports, the basin is assumed to have a zonal width of  $64^\circ$ . The value of  $\Psi^+ + \Psi^-$  is zero when the MOC is symmetric with respect to the equator.

We are interested in transitions from the top branch to the bottom branch of the bifurcation diagram. We will investigate transitions at two parameter values  $\beta_a = -0.2162$  and  $\beta_b = -0.1982$ . The transitions will be from point  $a$  to  $a'$  and from point  $b$  to  $b'$ . The pattern of the asymmetric overturning streamfunction at location  $b$  with sinking in the northern part of the basin is shown in Figure 6.1b, as well as the overturning streamfunction at location  $b'$  with sinking in the southern part of the basin. The saddle-node bifurcations are located at  $\beta = -0.2320$  and  $\beta = 0.3371$ .

## 6.2.2 Stochastic freshwater forcing

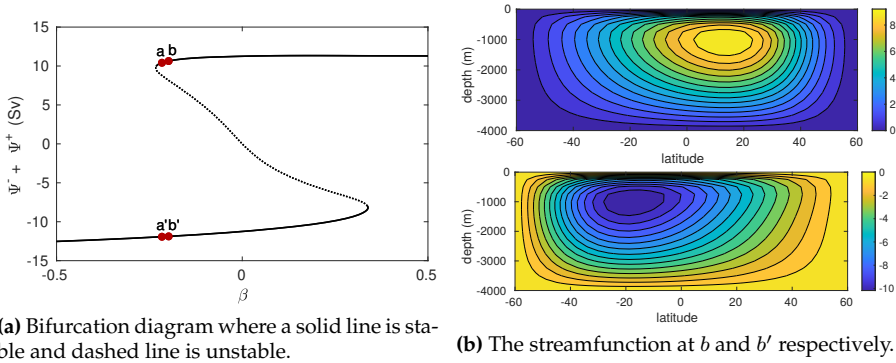
For the stochastic model, we use the same freshwater forcing as given in Section 2.5.3. However, because we use a different deterministic forcing, we can not simply describe  $F_s$  in terms of  $\bar{F}_s$  as given in Section 6.2.1. The more precise description is given by

$$F_s(\theta, t) = (1 + \sigma \zeta(\theta, t)) \mu F_0 \frac{\cos(\pi\theta/\theta_n)}{\cos(\theta)} + \beta F_p(\theta), \quad (6.3)$$

where  $\zeta(\theta, t)$  is as described in Section 2.5.3.

## 6.2.3 Reaction coordinate

Since  $p$  and  $w$  are determined by the algebraic constraints up to a constant, and therefore do not have a unique solution, computing the reaction coordinate



**Figure 6.1:** (a) Bifurcation diagram of the deterministic 2D MOC model, with the forcing as in (6.1). (b) Streamfunction pattern at  $b$  and  $b'$  respectively.

by using the norm of the state is not possible. We instead want to define the reaction coordinate in such a way that the distance between the meridional streamfunctions goes to zero, and therefore we define a norm on the  $v$ -part of the state. Combining this with (5.8), this gives us a reaction coordinate of the form

$$\phi(\mathbf{x}) = \eta - \eta e^{-8\|\mathbf{x} - \bar{\mathbf{x}}_A\|_v^2} + (1 - \eta)e^{-8\|\mathbf{x} - \bar{\mathbf{x}}_B\|_v^2},$$

where  $\|\cdot\|_v$  denotes the 2-norm on only the  $v$ -velocities, and  $\eta = \|\mathbf{x}_C - \bar{\mathbf{x}}_A\|_v / \|\mathbf{x}_B - \bar{\mathbf{x}}_A\|_v$  is the normalized difference between the unstable steady state  $\mathbf{x}_C$  and the stable steady state  $\mathbf{x}_A$ . We assume that a trajectory has reached  $B$  if  $\phi(\mathbf{x}) > z_{\max} = 0.75$ , since  $\mathbf{x}_C$  is really close to  $\mathbf{x}_A$ .

## Results 6.3

In this section we report on the transition probabilities that we find and compare the results from TAMS with the results from its projected variant.

Computations were performed on Peregrine, the HPC cluster of the University of Groningen. Peregrine has nodes with 2 Intel Xeon E5 2680v3 CPUs (24 cores at 2.5GHz) and each node has 128 GB of memory. For every run of TAMS, four cores were used to allow for efficient parallelization. Due to the problem size and the way the vertical mixing is implemented, it did not make sense to use more than four cores.

For our experiments, we investigated the probability of a transition within 2000 years. We ran TAMS 100 times with 300 samples. The stochastic theta method was used with  $\theta = 0.5$  and a time step of 2 years. For projected TAMS we used the approximate solution of the generalized Lyapunov equation with a tolerance of the relative residual of  $10^{-8}$ . The results are shown in Table 6.1.

Method	$\beta$	$\beta + 0.2320$	Mean	Variance
TAMS	-0.2162	0.0158	0.0495	$1.48 \cdot 10^{-4}$
Projected TAMS	-0.2162	0.0158	0.0496	$1.47 \cdot 10^{-4}$
TAMS	-0.1982	0.0338	$2.52 \cdot 10^{-7}$	$2.16 \cdot 10^{-13}$
Projected TAMS	-0.1982	0.0338	$3.36 \cdot 10^{-7}$	$5.35 \cdot 10^{-13}$

**Table 6.1:** Mean and variance of the probability of a transition within  $T = 2000$  yr with 100 runs of TAMS and the projected variant with 300 samples.  $\beta + 0.2320$  is the distance to the bifurcation point, which is located at  $\beta = -0.2320$ .

Here we see that a small step away from the bifurcation means a drastic decrease of the transition probability. Due to the convergence behavior of the relative error, the variance is relatively large for the rightmost point. We are only interested in the order of magnitude of the probabilities, however, and in the comparison between the two methods. To further decrease the variance in the future, we have to find a better reaction coordinate.

We notice a slight difference in the probability and an increased variance for projected TAMS at point  $b$ , which is the furthest away from the bifurcation. This can be explained by the fact that the probability is close to the tolerance of the Lyapunov solver. This may be solved by using a smaller tolerance for the Lyapunov solver, at the cost of more computational time and more memory usage. For the point that is closest to the bifurcation, the probabilities and variances are nearly the same.

In Table 6.2 we show the average computational time and memory usage of each time step in TAMS. We see that the projected method is approximately 30% faster, but more interestingly, uses only about 4% of the memory.

	$t(s)$	Mem (kB)
Normal time step	0.452	96
Projected time step	0.311	4

**Table 6.2:** Time and memory usage per time step in TAMS at point  $b$ .

To investigate if the time step of 2 years was small enough, we also ran the experiment at a time step of 1 year, of which the results are shown in Table 6.3. We see that the results are very similar, which confirms that a time step of 2 years is sufficiently small.

## 6.4 Summary and Discussion

Based on the results that were obtained in the previous chapter, we decided to use TAMS to compute transition probabilities in the idealized 2D MOC model. We showed that this is indeed possible, and observed that the transi-

Time step	$\beta$	Mean	Variance
2 yr	-0.1982	0.0495	$1.48 \cdot 10^{-4}$
1 yr	-0.1982	0.0493	$1.45 \cdot 10^{-4}$

**Table 6.3:** Different time steps for  $T = 2000$  yr with 100 runs of TAMS with 300 samples for time steps of 1 and 2 years.

tion probability decreased rapidly when moving away from the saddle-node bifurcation.

The computation of these transition probabilities is very expensive in terms of both time and memory usage. To be able to also compute transition probabilities on more realistic models, or in case the transition probability is small, the computational cost has to be reduced. The proposed projected time stepping method was shown to be 30% faster, and uses 96% less memory. We applied the projected time stepping method in TAMS, but the method may also be used in other methods for investigating rare events such as AMS and GPA.

In the future, it would be interesting to investigate the possibility of using a combination of most probable transition paths as computed in [Laurie and Bouchet \(2015\)](#) and a resampling method as used in AMS or GPA to obtain the transition probability in an even more efficient way. Such a method is very likely to still include stochastic time stepping, which means it can still be combined with the projected time stepping method described in this chapter.

Something that also needs further investigation is the reaction coordinate. In this thesis, we just chose a reaction coordinate that works, but it may be possible to find a reaction coordinate that works better. It would especially be nice if we can find a reaction coordinate that works well for a wide variety of high-dimensional systems.



