Model Order Reduction for Large Scale Wave Propagation

Victor Pereyra*

There are many applications in which it is necessary to solve large scale parametrized wave propagation problems repeatedly. This is still quite a challenging task, even with the largest available computer clusters. In this paper we will discuss the application of Model Order Reduction (MOR) to problems in seismic petroleum exploration, with the aim of diminishing the necessary computing time and storage by a significant factor. We consider POD and some variants. POD is a Model Order Reduction technique that uses snapshots of a few simulations in order to quickly compute related problems with similar accuracy. The method of lines via a Petrov-Galerkin approximation that uses the snapshots as basis functions without orthogonalization is the considered approach. The order reduction comes from projecting the wave equation discretized in space to the subspace spanned by the snapshots. This has been shown earlier to work well in two dimensions. The challenge in three dimensions comes from the size of the spatial meshes required and the fact that the method usually requires a number of snapshots that do not fit in fast memory, even for current high end multicore machines. Parallelization is not an option since it is already used for other aspects of this massive problem. We use several techniques to overcome these difficulties and show in some realistic large scale examples that they do work.

*Energy Resources Engineering, Stanford University, California.

1 Introduction

Seismic exploration is one of many approaches used to obtain information of the underground to locate and extract oil and gas. The seismic method is used inland and off-shore and consists of introducing man-made sources of energy that travels through the earth. The back scatter is recorded on sensors called geophones or hydrophones. It is this time data that is used to produce depth maps of the earth interior through seismic data processing. With the advent of powerful large scale clusters of computers it has become possible to simulate full wave propagation at various levels of complexity (acoustic, elastic, anisotropic) in order to migrate the data from time to depth to produce three-dimensional maps of the material properties that affect wave propagation, i.e.: density, wave speeds and anisotropic parameters.

A 3D seismic survey consists of many shots, i.e., activation of energy sources, say explosions, vibrator trucks or air-guns and corresponding receiver arrays. Usually, a simple rectangular geometry with equally spaced sources and receivers is used and only one source is activated at a time with a recording time of a few seconds, that depends on the depth one wants to image and the average velocity of the involved rock formations. Because of the complexity and the large scale of the 3D meshes involved the simulation of the time domain wave equation is performed using explicit methods on regular meshes, on a box that contains the region of interest. A typical 3D survey may have thousands of shots and for each shot thousands of receivers will be activated to record the back scatter, generating enormous data sets that have to be processed.

Repeated wave equation simulation is used for many different facets of this work, from survey planning and illumination studies, to imaging, migration from time to depth, reverse time migration, seismic tomography and quality control.

In recent work we have shown that model order reduction (MOR) is capable of reducing the computational cost of High Fidelity full acoustic wave simulation for shot interpolation and extrapolation in two-dimensions [4, 6, 8], when many problems with different source positions (shots) need to be simulated in a seismic survey or for seismic imaging. The requirement there is accuracy (albeit low, say 3 significant figures) and thus we cannot stray too far from the sources that produced the snapshots for MOR. We have shown that extrapolating one shot in each direction from the basis shot or interpolating two interior shots from two end shots works well when the data thus generated is used in a simulation process [8]. Taking into account the overhead, this doubles the speed at which we can perform this task, where many thousands of shots and corresponding simulations are required.

In [7] we discussed the three-dimensional problem in some detail and showed some preliminary results. In this paper we explore a number of avenues to solve larger problems competitively using some new variants of MOR. We also consider higher source frequencies and longer integration intervals.

For the application of interest, the integration domain is a half space that needs to be artificially limited on five sides, where absorbing boundary conditions are imposed. Thus, the geometry is very simple (a box), although it would be of interest to have topography, i.e., a non-planar surface as the top boundary. For this type of large wave propagation problems it is now routine to use explicit high order methods on uniform meshes, since they are the most efficient and simple ones available.

The acoustic wave equation in three dimensions with a forcing term and absorbing boundary conditions can be writen as:

$$w_{tt} = v^2(x,y,z)\triangle w + bu(t) - 2\epsilon(x,y,z)w_t - \epsilon^2(x,y,z)w,$$

where v is the velocity of propagation and the function ϵ decays rapidly away from the artificial boundaries. First, this equation is discretized in space on a mesh of size $n = nx \times ny \times nz$, and $k << n$ snapshots are collected by running one or several High Fidelity simulations. The snapshots are composed of values of the field variables $w(x,y,x,t)$ at the points of the spatial mesh for selected times, ordered in a vector with indices running first in the z direction and then in the y and x ones. They are written in this vector form as columns of an $n \times k$ matrix S.

An important new contribution consists of performing MOR using only a limited number of rows of the snapshot matrix (i.e., n_r rows associated with spatial mesh points selected through a $n_r \times n$ matrix C) and finally we show that we can gainfully use snapshots from selected simulations by only integrating part of the total time.

The reduction to a lower order system can be attained using directly the matrix of snapshots without orthogonalization, as we explain in the section on Oblique Projection, where a well conditioned basis is created using a progressive adaptive QR algorithm in reduced row space. We give numerical results for the oblique projection method for large problems in 3D with sources with up to 10 Hz central frequency and 4" of integration time.

Although most of the components of these approaches are not new, it is startling to find that very little work has been done to combine and apply them to the seismic exploration area, one of the industrial computations that uses a large number of flops on a regular basis. One of the main recent contributions corresponds to using MOR in the time domain for electromagnetic wave propagation (Maxwell equations) and in the frequency domain for the acoustic wave equation (Helmholtz equation), where calculations are shown for frequencies of up to 7.5 Hz [2].

2 MOR without orthogonality: Oblique Projections

We consider now the use of the snapshots directly in MOR, **without obtaining an orthogonal basis**, as is done in the classical approach. We assume then that $w(t) = Sa(t)$, where is the matrix of snapshots, and replace this expression in the original wave equation:

$$Sa_{tt} = ASa + \mathbf{b}u(t) - 2D(\epsilon)Sa_t. \tag{2.1}$$

Multiplying by S^T we get:

$$(S^TS)a_{tt} = S^T[ASa + \mathbf{b}u(t) - 2D(\epsilon)Sa_t].$$

The matrix coefficient of the first term in the right-hand-side can be obtained by solving the matrix least squares problem:

$$SX = AS,$$

where X is a $k \times k$ matrix.

3 MOR with a limited number of rows

We want to explore the possibility of only using a limited number of rows in 2.1, which, if possible and accurate, will go a long way to make this approach practical. For this we introduce:

$$r = Cw,$$

where the matrix C, $n_r \times n$, selects the elements of w that correspond to $n_r < n$ spatial positions (we assume that the source and receiver positions are included in that set) and packs them on a vector r of size n_r. Now, if we multiply equation 2.1 by C, the reduction proceeds in the usual fashion if we assume $w = Sa$:

$$CSa_{tt} = CASa + C\mathbf{b}u(t) - 2CD(\epsilon)Sa_t.$$

This step essentially selects the n_r equations corresponding to the desired output positions. In the finite element literature this is known as a Petrov-Galergin method [1].

Observe that the product CAS with the full matrix S can be performed column-wise in the High Fidelity (HF) code, as the snapshots are generated and then only n_r rows need to be saved to be read by the MOR code. The only proviso here is that the adaptive selection of a well-conditioned basis needs to be done before calculating this product and saving the snapshots. In fact, since even in the 3D case the sizes will be moderate if n_r is sufficiently small, we can go back to using adaptive QR to obtain an orthogonal basis U in the reduced space. In other words, we can calculate $CS = QR$, where Q is orthonormal and R is upper triangular and well conditioned and from now on S is a selected subset of all the snapshots inspected.

The other two terms in the right-hand-side only require the compacted n_r rows and therefore they can be calculated in the MOR code. This eliminates totally the large dimension n and should help to make possible running the MOR algorithm within the memory limitations that we have. Replacing we get:

$$QRa_{tt} = CASa + C\mathbf{b}u(t) - 2CD(\epsilon)Sa_t.$$

Finally:

$$a_{tt} = R^{-1}Q^T CASa + R^{-1}Q^T C\mathbf{b}u(t) - 2R^{-1}Q^T CD(\epsilon)Sa_t$$

and

$$w = CSa,$$

where R^{-1} stands for solving an upper triangular system of linear equations. Observe that no abbreviated matrix multiplications are necessary, so in principle, there are no additional approximations. We select the snapshots subject to the diagonal elements of R to be above a given threshold, in order to improve the conditioning of the base.

Possible ways to select C, i.e., the desired rows are:

1. Random.
2. A smaller box around the sources and receivers.
3. A coarser mesh.

We show in the computational section results for option 2. in 2D and for option 3. in 3D. Although 2. is attractive and easy to implement it does not work well in 3D since too much accuracy was lost in the examples we tried. Perhaps in larger models it would make sense to consider a sub-model around the source and receivers that contains all the arrivals for the integration time of interest.

4 Full survey simulation

A full survey simulation involving many thousands of shots will be partitioned in small subsets of the sort that we consider below in order to use MOR. The easiest implementation is one in which the subsets are adjacent but do not overlap. In this case they can be calculated in parallel and no communication will be necessary.

If we allow overlaps, considerable savings in computing time can be effected at the cost of programming complexity, increased storage and read/writes between fast memory and auxiliary storage. As usual these costs need to be balanced in order to attain a positive gain. We explain next a possible strategy that extends to a 2D array of sources what we did for a line of sources in [8]. In the 2D modeling case we needed only to save the snapshots corresponding to the far end of the segment in order to reuse them for the next segment. Now we have two directions to attend to and things are more complex.

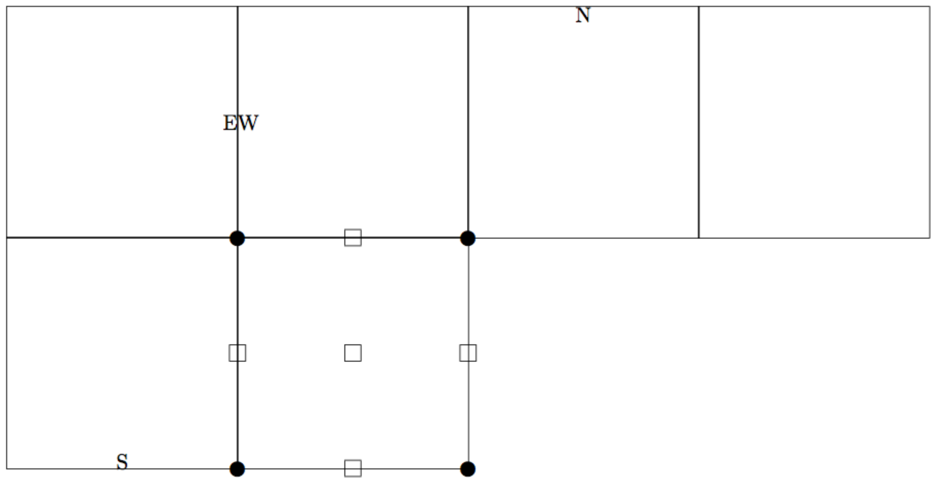

Figure 4.1: Schematic diagram of a full survey simulation

Algorithm

Collect the snapshots associated with each source in a single binary file with a header indicating the model name, shot number, length and number of snapshots. Assuming that we have a rectangle of subsets that form the whole shot array, the suggested algorithm proceeds as follows:

Start with the first row. After the first block is computed save the South (S) and East (E) snapshots.

For the next blocks in row one, re-use the E side of the previous block as the West (W) side of the new block, replace the old E side by the new one and save the new S side.

From block row 2 on, re-use S side of the block above and replace it by the new S side. Similarly with the E side (after the first block). Once a row is finished go to 3 until the whole survey has been simulated. For an interior 5×5 block as exemplified above, instead of 5 full integrations (assuming 1/10 of time steps for every other shot in the interior) we would have only 1.333 integrations for a 3.75× save. For our example above that would imply a total speedup factor of approximately 12.5!

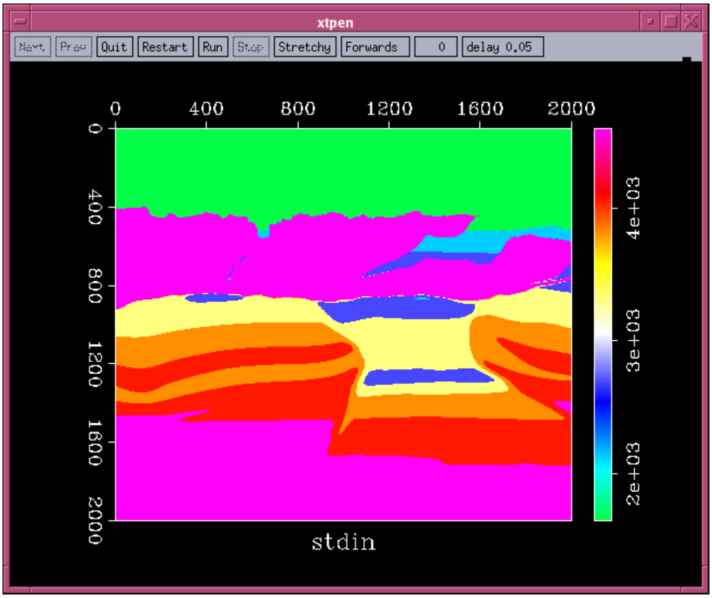

Figure 5.1: Model vel4pc velocity

Of course, for an specific problem one would also have to consider the possible disk access penalties. In the diagram of Figure 4.1 we present an schematic diagram of the geometry used.

If a distributed computation is desired the most effective way would be to use a frontal approach, in which a row and a column of blocks are calculated in parallel and that provides some of the redundant information for the next row and column.

5 A large scale 2D numerical example

We consider now *vel4pc*, a 2001×2001 $2D$ model of complex geology. See Figure 5.1 for a depiction of the 2D velocity mesh.

The number of time step stations (where we potentially inspect snapshots) is 400, with $dt = 0.005''$, for a total integration interval of 2" and a 5 Hz Ricker wavelet as the source. We consider 10 shots spaced by 6.25 units and extract a trace (i.e., time history) at $[900, 5]$ for comparison of the HF and MOR codes. We exercise the options of using less than the total integration interval in selected snapshots for the code

Shot	1	2	3	4	5
Error	0.00017	0.00053	0.00091	0.0012	0.0013
Shot	6	7	8	9	10
Error	0.0013	0.0012	0.00095	0.00057	0.00025

Table 1: Maximum absolute errors for the control trace

Method	Time (sec.)
HF	249579
HFS	48510
MOR	1168
Ratio	5.0

Table 2: Performance

HFS, namely: for the end shots, S_1 and S_{10} we use the first 300 time stations, while for the inner shots we only inspect the first 30 ones, for a total cost of about 2.1 full integrations. With a threshold at the end points of $\epsilon = 0.075$ and $\epsilon_i = 0.9^{i-1} * \epsilon,\ i = 2, ..., 9$ for the interior ones, a well-conditioned basis of 164 snapshots is selected by the adaptive QR algorithm applied to CS. We also consider the matrix C so that only 200100 consecutive points are used in HFS and MOR, starting at 1000500. In Table 1 we list the maximum absolute errors for the difference between the HF and MOR traces and in Figure 5.2 we show cross-plots for the best and worst cases ($S_{1,}S_5$).

Observe that for this example there is no action after time station 300. If there were significant signal after time station 300 then we could not ignore this final interval when seeking useful snapshots from the end sources, as we have seen in other examples. Again we see that phases are very good, while there is a miss in amplitude matching, which is visible at this scale. We observe in Table 2 that a speed up by a factor of 5 has been achieved by been able to use a significantly larger number of shots to share the overhead. This should be even more pronounced in 3D. We have used no parallelization (one thread only for the three codes).

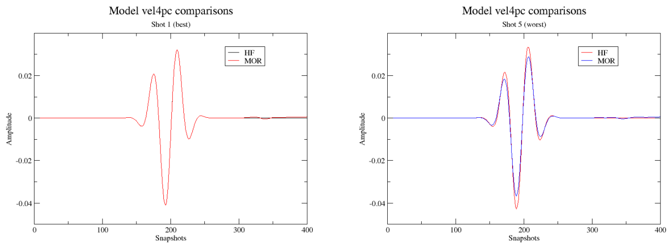

Figure 5.2: Cross-plot of best and worst shot

6 A 3D version

We consider a subset of the previous model, extended laterally, for $301 \times 301 \times 201 = 18,210,801$ mesh points and call it BP125. The time increments for the snapshot inspection stations is $dt = .02$", for an integration interval of 4". This is double the integration interval of the previous example and it shows a number of later arrival diffraction from the rugose top of the salt. In order to make it more challenging, we also double the frequency of the source to 10 Hz. We consider first 5 shots on a line run.

We use a variable threshold to select the snapshots. The initial value is $thresh = 1.0$. For each new shot this is replaced by $thresh = 0.7 * thresh$, so that there is a better chance of selecting snaps from later shots. In the same spirit, the threshold within a shot is decreased linearly from its original value to $thresh_{low} = 0.1$, again favoring the inclusion of later time snapshots. For the first and last shot we do a full integration while for the even interior shots we only integrate for the first 20 time stations, while for the central shot we use the first 40 time stations.

We also test the use of a C that selects a coarser mesh, by taking every third point in each direction, for a reduction of the size of the problem by a factor of 27, $n_r = 674,480$. This number includes the 5 sources and 1 receiver that are not in the original coarser mesh.

The performance table compares the extrapolated results for a 5×5 shot molecule with no help from adjacent ones and then for the interior ones, according to the strategy outlined in Section 4. We also show a cross-plot of the response at a particular point for both HF and MOR for two shots. For shot 1 there is a very good agreement, while for

Shot	1	2	3	4	5
Error	0.00062	0.00099	0.0015	0.002	0.0025
kc	135	140	159	163	243

Table 3: Maximum absolute errors for the control trace. 3D problem.

Method	Time 1 shot (sec.)	25 shots	25 shots with memory
HF	22702	567550	454040
HFS	25120	175840	45216
MOR	568	14200	9088
Ratio	-	2.98	8.36

Table 4: Performance

shot 5 we see some discrepancies, most noticeably the low amplitude oscillations after the main arrival starting at time station 20.

6.1 Conclusions

We have described an important set of problems pertaining to oil exploration by seismic methods. It is now routine to use full wave equation simulations to try to image the earth interior. Many such related simulations are required and we have discussed the application of Model Order Reduction in an attempt to decrease the computational effort by a significant factor. We have already shown in 2D that these techniques are effective, but in 3D the problems are very large and tax the

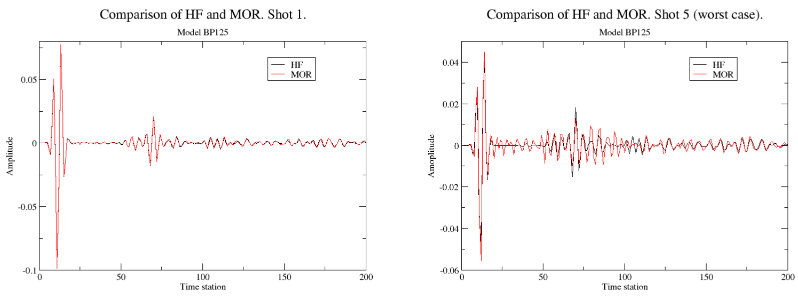

Figure 6.1: HF vs MOR for one trace, shot 1 (best) and 5 (worse).

largest memories available, especially for MOR based on snapshots. Thus we have reviewed a number of techniques that potentially can help in reducing this burden and we have shown numerical results for some of them that indicate that one can speed up these large calculations by near an order of magnitude. The large 3D results have been obtained by combining slanted projections, deterministic reduction and compactification of the number of rows (spatial mesh positions) of the problem and adaptive QR to select a well conditioned basis of snapshots for the reduced problem.

References

[1] K. Carlberg, C. Bou-Mosleh, and C. Farhat, *Efficient nonlinear model reduction via a least-squares Petrov-Galerkin projection and compressive tensor approximations.* Int. J. Numer. Meth. Engng. **86**:1-25 (2011).

[2] V. Druskin, R. Remis and M. Zaslavsky, *An extended Krylov subspace model-order reduction technique to simulate wave propagation in unbounded domains.* J. of Comp. Physics **272**:608-618 (2014).

[3] P. C. Hansen, V. Pereyra and G. Scherer, Least Squares Data Fitting with Applications. Johns Hopkins University Press, Baltimore (2013).

[4] B. Kaelin and V. Pereyra, *Fast wave propagation by model order reduction.* ETNA **30**:406-419 (2008).

[5] H. B. Keller and V. Pereyra, *Symbolic generation of finite difference formulae.* Math. Comp. **32**:955-971 (1978).

[6] V. Pereyra, *Wave equation simulation in two-dimensions using a compressed modeler.* American Journal of Computational Mathematics **3**:231-241 (2013).

[7] V. Pereyra, *Model order reduction with oblique projections for large scale wave propagation.* Journal of Computational and Applied Mathematics, **295:**103-114 (2016).

[8] Chunling Wu, Dimitri Bevc and V. Pereyra, *Model order reduction for efficient seismic modeling.* SEG Annual Meeting (2013).