Fast divergence-conforming reduced basis methods for stationary and transient flow problems

E Fonn1, H van Brummelen2, T Kvamsdal1,3 and A Rasheed1,4
1 Department of Applied Mathematics and Cybernetics, SINTEF Digital
2 Department of Mechanical Engineering, Eindhoven University of Technology
3 Department of Mathematical Sciences, Norwegian University of Science and Technology
4 Department of Engineering Cybernetics, Norwegian University of Science and Technology
E-mail: eivind.fonn@sintef.no

Abstract. Reduced basis methods (RB methods or RBMs) form one of the most promising techniques to deliver numerical solutions of parametrized PDEs in real-time with reasonable accuracy [1]. For the Navier-Stokes equation, RBMs based on stable velocity-pressure spaces do not generally inherit the stability of the high-fidelity method. Common techniques for working around this problem (e.g. [2]) have the effect of deteriorating the performance of the RBM in the performance-critical online stage.

We show how divergence-free reduced formulations eliminates this problem, producing RBMs that are faster by an order of magnitude or more in the online stage. This is most easily achieved using divergence-conforming compatible B-spline bases, using a transformation that can maintain the divergence-free property under variable geometries. See [3] for more details.

We also demonstrate the flexibility of RBMs for non-stationary flow problems using a problem with two stages: an initial, finite transient stage where the flow pattern settles from the initial data, followed by a terminal and infinite oscillatory stage characterized by vortex shedding. We show how an RBM whose data is only sourced from the terminal stage nevertheless can produce solutions that pass through the initial stage without critical problems (e.g. crashing, diverging or blowing up).

1. Introduction

Conventional methods for simulating partial differential equations include well-established techniques such as Finite Volume Methods (FVM), Finite Difference Methods (FDM) and Finite Element Methods (FEM). Common to all of these methods is the large number of degrees of freedom that is typically required to accurately model a physical system, often numbering in the millions or billions. Given the good and well-established approximation properties of FVM, FDM and in particular FEM, such models are usually classified as high-fidelity models. Problems of this size are not generally possible to solve in practical timeframes except on specialized hardware, and even then they may require several days of computing time. This prohibits the use of high-fidelity models in time-critical on-site analyses, e.g. for developing predictive digital twins [4].

This computational complexity is also at odds with the increasing demand for real-time low-cost models for the repetitive solution of physical models in many-query scenarios. This is particularly relevant in optimization, control systems, inverse and inference problems and
uncertainty quantification. Common for many of these applications is that the model in question is parametrized by a suitably small number of input parameters. It is often required to be able to provide solutions in the sub-second regime.

Reduced-order modeling (ROM) provides a paradigm to address the aforementioned challenges. ROM is a rapidly developing field [5]. The general aim of ROM is to replace the original model with a reduced model of very modest computational complexity. Within this general framework of ideas, one of the most promising is that of reduced basis methods (RBM). This method dates back to the 1980s with work from [6, 7, 8, 9, 10, 11, 12]. Excellent modern introductions can be found in [1, 13].

The fundamental concept of RB methods is to formulate the problem on a function space with very low dimension, where the basis functions are tailored to the solution of the model in the parameter regime of interest. In comparison, while e.g. FEM can boast well-established asymptotic approximation properties, the approximation power per degree-of-freedom is clearly limited. RB methods seek to construct a basis with optimal approximation properties for the given class of solutions, under the premise that the cost of constructing the basis is inconsequential. This premise arises because of the offline/online division. The offline stage, where the basis is constructed, is run only once, the result of which is a RBM which can be used in the online stage, to be run for each parameter query.

This paper is concerned with the application of a divergence-free RBM for Navier-Stokes flow problems. This method is detailed in [3] for stationary flow problems, where significant performance improvements were found relative to a traditional RBM with velocity, pressure and stabilizing ‘supremizer’ modes, as in [2]. Some of the theory and results from [3] is repeated here.

Further, we aim to investigate the applicability of RBMs to non-stationary flow problems characterized by multiple ‘stages’: a flow problem with a significant but finite ‘ramp up’ stage followed by a non-terminating flow pattern characterized by vortex shedding. In this context we envision that the vortex shedding stage is of primary interest, and it is therefore desirable to avoid including degrees-of-freedom which are sourced from the ramping-up stage. From an approximation utility point of view, such degrees-of-freedom are costly and not beneficial. We will demonstrate that an RBM whose solution space has been sourced only from the vortex shedding stage is able to produce a time-stepping sequence that can robustly reach this stage.

2. Parametrized Navier-Stokes equations

We consider the non-stationary Navier-Stokes equations,

\[ \partial_t \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p = f \quad \text{in } [0, T] \times \Omega, \]

\[ \nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega, \]

\[ \mathbf{u} = \mathbf{g} \quad \text{on } \Gamma_D, |\Gamma_D| > 0 \]

\[ -p \mathbf{n} + \nu (\nabla \mathbf{u}) \mathbf{n} = h \quad \text{on } \Gamma_N, \]

\[ \mathbf{u}(t = 0) = \mathbf{u}_0 \quad \text{on } \Omega. \]

where \( \nu \) is the viscosity, \( \mathbf{u}, p \) are the unknown velocity and pressure, \( f, g, h, \mathbf{u}_0 \) correspond to exogenous and initial data, \( \Omega \subset \mathbb{R}^d \) is the domain of interest with boundary \( \partial \Omega = \Gamma_D \cup \Gamma_N, \Gamma_D \cap \Gamma_N = \emptyset \), \( \mathbf{n} \) denotes the external unit normal vector, and \( T \) is the upper boundary in time.

The weak Galerkin formulation of the problem is to find \((\mathbf{u}, p)\) members of suitable function spaces, such that for all \((\mathbf{w}, q)\), also members of suitable function spaces, it holds that

\[ m(\dot{\mathbf{u}}, \mathbf{w}) + a(\mathbf{u}, \mathbf{w}) + c(\mathbf{u}, \mathbf{u}, \mathbf{w}) + b(p, \mathbf{w}) = d(w), \]

\[ b(q, \mathbf{u}) = 0, \]
where the linear, bilinear and trilinear forms \(m, a, b, c, d\) are defined as

\[
m(u, w) = \int_{\Omega} u \cdot w, \\
a(u, w) = \nu \int_{\Omega} \nabla u : \nabla w, \\
b(p, w) = -\int_{\Omega} p \nabla \cdot w, \\
c(u, v, w) = \int_{\Omega} (u \cdot \nabla) v \cdot w, \\
d(w) = \int_{\Gamma_N} \mathbf{h} \cdot w + \int_{\Omega} \mathbf{f} \cdot w.
\]

(8a) \hspace{1cm} (8b) \hspace{1cm} (8c) \hspace{1cm} (8d) \hspace{1cm} (8e)

In the following it will be assumed that all function spaces involved in this formulation are closed under linear combinations. This is generally not the case, unless \(g \equiv 0\), but the process for transforming a non-homogeneous model into a homogenous one by the use of a lift function is well known, see e.g. [3].

We now consider the case when the problem (1)–(5) depends on a number of parameters. We will denote by \(\mathcal{P}\) the parameter space, and by \(\mu\) any given element of \(\mathcal{P}\).

The effect of varying parameters influences not only the solutions \(u\) and \(p\), but also the weak forms (8a)–(8e) directly. This is most obviously the case for example with physical parameters such as the viscosity \(\nu\) in (8b), data parameters such as \(g, h\) (the former entering (8a)–(8e) via a parameter-dependent lift function), but one may also consider geometric parameters, where the domain \(\Omega\) itself may be transformed.

For geometric parameters, it is necessary to define a reference domain \(\hat{\Omega}\), a mapping \(\chi_\mu : \Omega(\mu) \to \hat{\Omega}\), and function space mappings for velocity and pressure,

\[
\pi^v_\mu : [H^1(\hat{\Omega})]^d \to [H^1(\Omega(\mu))]^d, \\
\pi^p_\mu : L^2(\hat{\Omega}) \to L^2(\Omega(\mu)).
\]

(9) \hspace{1cm} (10)

Thus, for example, (8b) can be properly parametrized as

\[
(a(\pi^v_\mu \tilde{u}, \pi^v_\mu \tilde{w}; \mu) = a(\pi^v_\mu u, \pi^v_\mu v; \mu)
\]

(11)

In this way, the entirety of (8a)–(8e) can be formulated as a parameter-dependent homogeneous problem on a fixed reference domain, involving fixed, function spaces for velocity and pressure which are closed under linear combinations. See [3] for further details.

3. Model order reduction

To apply the RB method, we first solve the high-fidelity Navier-Stokes problem for a suitably large and varied choice of parameter instances \(\mu\). The resulting ensemble is then compressed using Proper Orthogonal Decomposition (POD) [14] into a reduced basis which can be represented as a tall matrix \(V\) of \(m\) columns, where each column represents one reduced basis function.

The reduced system of equations can then be formulated. Given a high-fidelity model represented as

\[
A(\mu)u(\mu) = f(\mu)
\]

(12)

where \(A(\mu)\) and \(f(\mu)\) are the discrete system matrix and right hand side arising from (6)–(7), we make the assumption that the solution coefficient vector \(u(\mu)\) can be written in terms of the columns of \(V\),

\[
u(\mathcal{U}) \approx V\mathbf{u}_r(\mu)
\]

(13)
where the vector $\mathbf{\tilde{u}}_r(\mu)$ is the coefficient vector in the reduced basis. Since (12) is now overdetermined, we can reduce the system from the left in a similar manner, effectively forming a reduced Galerkin formulation.

$$V^\top A(\mu)V\mathbf{u}_r(\mu) = V^\top f(\mu).$$

(14)

For a modest number of reduced basis functions $m$, this system is small and quickly solvable. However, unless certain assumptions on $A$ and $f$ are met, it may not be possible to quickly assemble it. We call this the assumption of affine representations: that $A$ and $f$ can be written in the form

$$A(\mu) = \sum_{i=1}^{M} \theta_i(\mu) A_i, \quad f(\mu) = \sum_{i=1}^{N} \xi_i(\mu) f_i.$$  

(15)

This allows us to compute all matrices $V^\top A_i V$ in the offline stage, whence

$$V^\top A(\mu)V^\top = \sum_{i=1}^{M} \theta_i(\mu) [V^\top A_i V]$$

is easily and quickly assembled in the online stage.

The assumption (15) is not always easily realized, especially when it comes to geometric parameters. In many cases it may be necessary to approximate it with various interpolation techniques, such as the Empirical Interpolation Method (EIM) [1]. For the model problems in [3] the affine representation was produced with explicitly truncated series expansions.

4. Divergence-free reduced basis methods

It can be readily observed from (6)–(7) and (8a)–(8e) that if the velocity solution and test spaces are fully divergence-free, the expressions involving the $b$-form vanish, and the continuity equation is trivially satisfied. This desirable state of affairs is not easily realized with high-fidelity methods, because generic divergence-free function spaces are difficult to make.

Isogeometric analysis (IGA) has recently come to the forefront in this field. While IGA velocity basis functions are also not in themselves divergence-free, the method, more easily than for classical finite element methods, allows the formulation of divergence-conforming discretizations, which produce strongly divergence-free solutions [15, 16, 17, 18, 19, 20, 21, 22].

In a classic RBM setting, each basis function is a linear combination of weakly divergence-free high-fidelity solutions, thus also weakly divergence-free. If, as with IGA, the high-fidelity method can produce strongly divergence-free solutions, the reduced basis functions will also be strongly divergence-free, so a velocity-only reduced formulation is possible. The pressure solution can be recovered, if necessary, using supremizers [2] as a stabilizing velocity test space. The details for this can be found in [3].

Divergence-free RBMs can be significantly faster than conventional RBMs, as we demonstrate in Section 5. The reason for this is that the velocity solution can be achieved with the solution of a linear system whose size is equal to the dimension $m$ of the reduced velocity basis. If required, the pressure can then be recovered through another Petrov-Galerkin linear system of size $m$ using supremizers as test functions. On the other hand, without divergence-free basis functions, the full $3m \times 3m$ system cannot be cleanly decoupled.

In case of parameter-dependent domains, it is necessary that (9) is divergence-conforming, i.e. it maps divergence-free functions in one geometry to divergence-free functions in the reference geometry. This is generally not the case for simple pullback transforms, but will be satisfied by e.g. the Piola transform [3, equation (64)].
5. Numerical example: Stationary 2D flow around NACA0015 airfoil

The following example considers two-dimensional stationary Navier-Stokes flow around a NACA0015 airfoil. Flow around such airfoils are relevant for harvesting wind energy, as cross sections of wind turbine blades typically are at least partly composed of NACA airfoils. High fidelity models using isogeometric finite elements have been developed in [23, 24].

The airfoil is suspended at the center of an “O”-mesh, and Uniform inflow velocity is applied as a Dirichlet condition at the left semicircle, \( g = (u_\infty, 0) \). The airfoil has varying angles of attack \( \varphi \), realized by rotating the entire mesh through an angle that varies with the distance from the center, in such a manner that the mesh deformation vanishes at the external boundary, as seen in Figure 1. This is done in order to maintain the parameter-independence of the Dirichlet and Neumann boundary sets \( \Gamma_D \) and \( \Gamma_N \).

The parameters of interest for the RBM study is \( \mu = (\varphi, u_\infty) \) with
\[
P = \{(\varphi, u_\infty) \mid \varphi \in [-35\degree, 35\degree], u_\infty \in [1 \text{ m/s}, 20 \text{ m/s}]\}
\]

The viscosity \( \nu \) was fixed at \( \nu = \frac{1}{6} \text{ m}^2/\text{s} \), and the airfoil has a chord length of 1 m, giving an approximate maximal Reynolds number of \( \text{Re} = 120 \).

Two discretizations were compared: a conventional Taylor-Hood method with quadratic velocities and linear pressures, and a divergence-conforming discretization with mixed quadratic-linear velocities and linear pressures. The latter method produces strongly divergence-free reduced velocity basis functions, enabling the aforementioned divergence-free RBMs. The derivation of affine representations for this geometry transformation is a nontrivial matter, out of scope for this paper. We refer to [3] where this is performed explicitly as a truncated Taylor expansion in \( \varphi \).

For both discretizations, an ensemble of 225 snapshot solutions were generated on the \( 15 \times 15 \) Gauss quadrature points in the parameter domain \( P \).\(^1\) RBMs were then generated with \( M = 10, 20, 30, 40, 50 \) basis functions for each of the reduced bases, giving a total of

\(^1\) The generation of an RBM using snapshots on a parameter space is theoretically equivalent to minimizing the integral of an error expression by substituting the integral with a quadrature expression, see [1]. For this reason, it is suitable to sample the parameter space at the Gauss points.
Figure 2. NACA0015 Airfoil: Measured mean relative error as a function of mean time usage (in seconds), for velocity ($H^1$-seminorm) and pressure ($L^2$-norm). The markers correspond to $m = 10, 20, 30, 40, 50$ degrees of freedom in velocity and pressure separately. The time reported includes both the velocity solution and the pressure recovery or reconstruction, as appropriate. At an error level of $\sim 1\%$ the divergence-free method has a speedup factor of $\sim 10 \times$. Bottom left is better.

$m = 30, \ldots , 150$ degrees of freedom in total, corresponding to $M$ each for velocity, pressure and supremizers, the latter of which play the role as a stabilizing space. The performance of the two discretizations was evaluated on a set of $15 \times 15$ uniformly spaced points in the parameter space, chosen so as to avoid the parameter values used to generate the ensemble.

For evaluating an RBM, the principal quantity of choice is the mean relative error, defined as the mean expected relative error between the high-fidelity solution and the reduced solution, as taken over a suitably dense sampling of the parameter space.

Figure 2 shows the true mean relative error as a function of online solver time. This reveals the divergence-free reduced basis method to be significantly faster for comparable accuracy, because it allows a velocity and a pressure solution to be found through one $m \times m$ linear system each, rather than a full $3m \times 3m$ linear system as with conventional methods. At an error level of roughly $1\%$ the divergence-free RBMs are already faster by a factor of about 10, which can be extremely valuable in the time-sensitive online stage.

Finally, Table 1 shows the mean wall-time taken for a high-fidelity or reduced solution for both methods. This again highlights the significant speed improvements achievable with RBMs, but also how quickly they can deteriorate as the reduced basis grows. The divergence-conforming method scales much better in terms of reduced degrees-of-freedom.

6. Numerical example: Transient 2D flow around a cylinder
The purpose of this section is to argue for the viability of limited reduced bases for non-stationary flow problems. The model problem is flow around a cylinder at $Re = 100$ without any parameter-dependence other than time. This is a setup that is known to produce vortex shedding, as seen
| $m$ | Taylor-Hood | Conforming |
|-----|-------------|------------|
|     | Hi-Fi | Time | Speedup | Time | Speedup |
| 10  | 39 s  | 110 s |
| 20  | 16 ms | 2400  | 10 ms  | 11000 |
| 30  | 79 ms | 490   | 17 ms  | 6500  |
| 40  | 290 ms| 130   | 32 ms  | 3400  |
| 50  | 750 ms| 52    | 61 ms  | 1800  |

Table 1. NACA0015 Airfoil: Mean timings and speedup factors for the Taylor-Hood and the divergence-conforming methods for different numbers of reduced degrees-of-freedom.

Figure 3. 2D cylinder flow: Vortex shedding in the wake of the cylinder at $Re = 100$. High-fidelity solution on the left, and an RBM with $m = 10$ degrees of freedom on the right. The coloring is according to flow speed magnitude.

Typically it takes several timesteps before a simulation reaches the stage where vortex shedding occurs continuously without significant variation over time scales other than the primary shedding frequency. We call this the terminal stage, and we propose that this is the stage of primary interest in many RBM applications. For example, in optimal design applications, one is generally interested in optimizing for the long term behaviour of a system. In contrast, the time steps before reaching the terminal stage, which we shall call the initial stage, are often of limited interest.

When designing an RBM for a non-stationary problem it is necessary that the reduced basis provides sufficient approximative power in every stage. If the solution at the end of the initial stage is poor, it is natural to imagine that no value can be found in the solution for the terminal stage. One would conclude, therefore, that a successful reduced basis must account for degrees-of-freedom from both stages. In this case, however, the vortex shedding behaviour develops naturally from almost any initial condition. We argue that, in the application context previously introduced, if an RBM can reach the terminal stage in any state whatsoever from which vortex shedding can develop, such a method is equally as powerful as an RBM with full approximative power for every stage.

In the following example, we performed a high-fidelity simulation around a cylinder of radius 1 m with an inflow velocity of $u_\infty = 1$ m/s at timesteps of $\Delta t = 0.5$ s. The viscosity
Figure 4. 2D cylinder flow: Long-term evolution of the drag force of limited RBMs in the initial stage, as compared to the high-fidelity method.

was fixed at $\nu = 1/100 \text{ m}^2/\text{s}$ giving a Reynolds number of $Re = 100$. After vortex shedding stabilized, we collected 500 snapshots. Thus, no snapshots were collected from the initial stage. As an initial condition, we used the stationary Stokes flow solution.

Figure 4 shows the behaviour of the drag force on the cylinder from four reduced models compared to the high-fidelity method. We can see that the RBMs behave very differently in the initial stage, and remain there for many more timesteps, but that they all reach the terminal stage in due time, and that the drag force as reported by the reduced methods correspond well to that reported by the high-fidelity method.

The timing results for this example are comparable to those in Table 1 per timestep, with the caveat that the RBMs take 4–6× as many timesteps to reach the terminal stage. The practical speedup factors are therefore about 600 for the Taylor-Hood method, and 2800 for the divergence-conforming method at $m = 10$. For higher numbers of degrees-of-freedom, the deterioration accelerates because of the tendency for larger RBMs to take longer to reach the terminal stage, as seen in Figure 4. It is therefore of particular importance that the divergence-conforming method has good performance characteristics for large $m$. It is also possible to use a smaller RBM for the initial stage, only expanding the solution space once the terminal stage is reached. Since reduced bases of different orders are nested, such a technique would not be difficult to implement.

7. Conclusions
We have demonstrated a significant advantage to using divergence-conforming high-fidelity methods to produce divergence-free reduced basis methods. The resulting RBMs are considerably faster ($\sim 10\times$) in the online stage than conventional RBMs with comparable accuracy ($\sim 1\%$ relative error). This is owing to the “decoupling” effect of divergence-free
basis functions, allowing separate solutions for velocity and pressure with two smaller linear systems, as opposed to a simultaneous larger system for both fields.

We have also shown that limited numbers of basis functions are suitable for non-stationary applications where the primary interest lies in the stable long-term behaviour of the system, as opposed to the initial stages. Snapshots can be sourced only from the terminal stage of interest without regard for the initial stages.
Acknowledgments
The authors acknowledge the financial support from the Norwegian Research Council and the industrial partners of OPWIND: Operational Control for Wind Power Plants (Grant No.: 268044/E20). https://www.sintef.no/en/projects/opwind/

References
[1] Quarteroni A, Manzoni A and Negri F 2016 Reduced basis methods for partial differential equations (Springer International Publishing)
[2] Ballarin F, Manzoni A, Quarteroni A and Rozza G 2015 International Journal for Numerical Methods in Engineering 102 1136–1161 ISSN 1097-0207
[3] Fonn E, van Brummelen H, Kvamsdal T and Rasheed A 2019 Computer Methods in Applied Mechanics and Engineering 346 486–512
[4] Rasheed A, San O and Kvamsdal T 2020 IEEE Access 8 21980–22012
[5] Bazaz M A, un Nabi M and Janardhanan S 2012 2012 IEEE International Conference on Signal Processing, Computing and Control pp 1–6
[6] Almroth B, Stern P and Brogan F 1978 Aiaa Journal 16 525–528
[7] Almroth B, Stelhin P and Brogan F 1981 22nd Structures, Structural Dynamics and Materials Conference p 575
[8] Nagy D A 1979 Computers & Structures 10 683–688
[9] Noor A K and Peters J M 1980 Aiaa journal 18 455–462
[10] Noor A K and Peters J M 1981 Computer Methods in Applied Mechanics and Engineering 29 271–295
[11] Noor A K 1981 Computational Methods in Nonlinear Structural and Solid Mechanics (Elsevier) pp 31–44
[12] Noor A K 1982 Computer methods in applied mechanics and engineering 34 955–985
[13] Haasdonk B 2017 Model reduction and approximation: theory and algorithms 15 65
[14] Chatterjee A 2000 Current science 808–817
[15] Buffa A, Sangalli G and Vázquez R 2010 Computer Methods in Applied Mechanics and Engineering 199 1143–1152
[16] Buffa A, De Falco C and Sangalli G 2011 International Journal for Numerical Methods in Fluids 65 1407–1422
[17] Evans J A and Hughes T J 2012 Computational Mechanics 50 667–674
[18] Evans J and Hughes T 2013 Mathematical Models and Methods in Applied Sciences 23 1421–1478 ISSN 0218-2025
[19] Evans J A and Hughes T J R 2013 Mathematical Models and Methods in Applied Sciences 23 671–741
[20] Evans J A and Hughes T J 2013 Journal of Computational Physics 241 141–167
[21] Johannessen K A, Kumar M and Kvamsdal T 2015 Computer Methods in Applied Mechanics and Engineering 293 38–70
[22] van Opstal T M, Yan J, Coley C, Evans J A, Kvamsdal T and Bazilevs Y 2017 Computer Methods in Applied Mechanics and Engineering 316 859–879
[23] Nordanger K, Holdahl R, Kvamsdal T, Kvarving A M and Rasheed A 2015 Computer Methods in Applied Mechanics and Engineering 290 183–208
[24] Nordanger K, Holdahl R, Kvarving A M, Rasheed A and Kvamsdal T 2015 Computer Methods in Applied Mechanics and Engineering 284 664–688