



Data Assimilation by Stochastic Ensemble Kalman Filtering to Enhance Turbulent Cardiovascular Flow Data From Under-Resolved Observations

Dario De Marinis^{1,2*} and Dominik Obrist¹

¹ ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland, ² Dipartimento di Meccanica, Matematica e Management and Centro di Eccellenza in Meccanica Computazionale, Politecnico di Bari, Bari, Italy

OPEN ACCESS

Edited by:

Zahra K. Motamed,
McMaster University, Canada

Reviewed by:

Humberto Godínez,
Los Alamos National Laboratory
(DOE), United States
Amirhossein Arzani,
Northern Arizona University,
United States
Pierre Sagaut,
Aix Marseille University, France

*Correspondence:

Dario De Marinis
dario.demarinis@poliba.it

Specialty section:

This article was submitted to
Cardiovascular Imaging,
a section of the journal
Frontiers in Cardiovascular Medicine

Received: 15 July 2021

Accepted: 06 October 2021

Published: 02 November 2021

Citation:

De Marinis D and Obrist D (2021) Data Assimilation by Stochastic Ensemble Kalman Filtering to Enhance Turbulent Cardiovascular Flow Data From Under-Resolved Observations. *Front. Cardiovasc. Med.* 8:742110. doi: 10.3389/fcvm.2021.742110

We propose a data assimilation methodology that can be used to enhance the spatial and temporal resolution of voxel-based data as it may be obtained from biomedical imaging modalities. It can be used to improve the assessment of turbulent blood flow in large vessels by combining observed data with a computational fluid dynamics solver. The methodology is based on a Stochastic Ensemble Kalman Filter (SEnKF) approach and geared toward pulsatile and turbulent flow configurations. We describe the observed flow fields by a mean value and its covariance. These flow fields are combined with forecasts obtained from a direct numerical simulation of the flow field. The method is validated against canonical pulsatile and turbulent flows. Finally, it is applied to a clinically relevant configuration, namely the flow downstream of a bioprosthetic valve in an aorta phantom. It is demonstrated how the 4D flow field obtained from experimental observations can be enhanced by the data assimilation algorithm. Results show that the presented method is promising for future use with *in vivo* data from 4D Flow Magnetic Resonance Imaging (4D Flow MRI). 4D Flow MRI returns spatially and temporally averaged flow fields that are limited by the spatial and the temporal resolution of the tool. These averaged flow fields and the associated uncertainty might be used as observation data in the context of the proposed methodology.

Keywords: data assimilation, Ensemble Kalman Filter, cardiovascular flow, turbulence, ensemble averaging

1. INTRODUCTION

The clinical relevance of 4D Flow Magnetic Resonance Imaging (4D Flow MRI) for quantifying pulsatile and turbulent blood flow in the ascending aorta is limited by spatial and temporal resolution which are in general insufficient for a precise assessment of flow related parameters such as turbulent kinetic energy (TKE), Reynolds shear stress (RSS) and wall shear stress. In recent years, models to overcome these limits have been proposed, e.g., to quantify TKE for assessing aortic stenosis severity (1).

Data Assimilation (DA) can help to enhance the quality of these parameters. In DA, sparse and noisy measurement data (observations) are combined with the forecast solution computed by a

numerical (forward) model in order to obtain an improved prediction of the true state of the system. A widely used and well-known technique for DA is the Kalman Filter (KF) proposed by Kalman (2). Computation of an appropriate filter is based on the uncertainties of the observations and of the forward model. These uncertainties can be described by covariances of specific variables that are typically not known *a priori*, such that the appropriate design of a filter requires modeling. The original KF can be applied to a wide range of applications if the physical system can be described by a linear model and data observed from measurements are affected by normally distributed noise. In this context, the forecast computed by the linear model is perturbed by the error of the model, and the forecast mean depends linearly only on the previous time step since the error has zero mean. The covariance of the error is not zero, therefore it also appears in the forecast covariance matrix computation.

In nonlinear systems, the forecast mean and covariance cannot be calculated directly from the previous time step anymore, because of the nonlinear nature of the model operator. This limitation led to enhancements of the basic KF theory. A first example is the Extended Kalman Filter (3) which linearizes the original nonlinear dynamics around the previous state estimates. The ensemble KF (EnKF) (4) replaces the forecast covariance matrix by a sample covariance and estimates the forecast mean and covariance from an ensemble of states of system, which represents the evolution of the state probability density function. This makes the EnKF an important DA tool for ensemble forecasting. If the observations can be interpreted as the result of an ensemble of samples, the noise covariance is replaced by the sample covariance which results in the Stochastic Ensemble Kalman Filter (SEnKF) (5, 6) also known as Ensemble Kalman Filter with perturbed observations.

Many examples of KF techniques applied to fluid dynamics problems have been reported in the literature. Hoepffner et al. (7) used a KF based on the linearized Navier–Stokes equations to reconstruct the relevant statistics of the initial conditions in transitional wall-bounded flow systems. The study has been extended to turbulent wall-bounded flows to estimate the mean turbulent flow profile in the near-wall region by using noisy observations on the wall (8) and to capture the turbulent flow state at the outer boundary of the buffer region of turbulent boundary layer by using an EnKF (9). Gu and Oliver (10) proposed an EnKF to investigate multiphase flows in porous media. Harlim and Majda (11) compare different methods for filtering sparsely observed turbulent geophysical flows in the atmosphere and ocean regimes. Suzuki (12) developed a hybrid unsteady-flow simulation technique combining particle tracking velocimetry (PTV) and direct numerical simulation (DNS) by using a reduced-order KF. Recently, a reduced-order model based on KF has been proposed for turbulent flow configurations showing a successful improvement of the prediction of turbulent features even when the observation is provided only in a limited region (13).

DA using methods different from KF have been also applied to fluid dynamics problems. A variational DA technique based on the minimization of the error between observations and numerical solution in the context of Reynolds Averaged

Navier–Stokes (RANS) equations has been applied successfully to reconstruct the mean flow field around a cylinder in laminar regime (14). The variational DA technique has been used to combine mean velocity from 2D PIV observations of flow over an idealized airfoil and a numerical solver for RANS Equation (15).

Both variational and KF-based DA techniques for computational fluid dynamics have been reviewed and compared for unsteady viscous flow applications (16). These techniques are powerful tools, because the coupling of observed data with computational models can remove errors which cannot be identified by using only one of these scientific methodologies alone. However, reliable tools for accurate prediction of complex flow configurations are still lacking.

In the field of cardiovascular flows, DA and Kalman filter have been used in order to improve the accuracy and reliability of physical modeling and to reduce the uncertainty due to the lack of information about boundary conditions, patient-specific geometries and blood viscosity. Gaidzik et al. (17) improved hemodynamic flow prediction by merging Phase-Contrast MRI data with CFD simulations for an idealized aneurysm model where well-controlled laminar flow can be obtained. Canuto et al. (18) implemented an EnKF for the purpose of estimating parameters in cardiovascular models, i.e., a fully zero-dimensional model of the right heart and pulmonary circulation and a coupled 0D–1D model of the lower leg, through the assimilation of clinical measurements of specific patients. DeVault et al. (19) proposed a model for the blood flow in a vital subnetwork of the cerebral vasculature, namely the Circle of Willis. In this model the parameters of the outflow conditions were calibrated using a subset of clinical measurements through EnKF techniques. Arnold et al. (20) used EnKF to estimate the inlet flow waveform in patient-specific arterial network models. Habibi et al. (21) used a reduced-order modeling Kalman filter to provide blood flow data that were more accurate than the computational and synthetic voxel-based experimental datasets with the aim of improving near-wall hemodynamics quantification. In their recent review, Arzani and Dawson (22) present and compare different variational and KF-based DA for modeling cardiovascular flows. Some other studies on merging CFD and 4D flow MRI data using data assimilation are reported here for completeness (23–27).

The aim of this paper is to propose a DA methodology based on the SEnKF approach that can be used to enhance the spatial and temporal resolution of voxel-based flow observations of turbulent pulsatile flow (as in 4D Flow MRI) to improve the assessment of turbulent blood flow in large vessels. To the best of the authors' knowledge, the current work seems to be the first study that tackles turbulence modeling in blood flow with KF-based DA techniques. We propose to consider voxel-based observed data as the result of a volumetric averaging of the true state over the voxel size. The associated sample covariance will take into account the presence of turbulence as well as the noise. The present problem of DA could also be addressed by the 4D-Var method (28, 29) or the Ensemble Kalman Smoother (30) which cope with problems presenting nonlinearity and strong sensitivity to initial conditions as in turbulent flows. Here, we only aim at finding the correct covariance matrix of these

fluctuations rather than trying to predict exact, instantaneous fluctuations. The sensitivity to initial conditions ensures that every simulated pulse results in a slightly different realization of the turbulent flow field. The average of this ensemble of flow realizations corresponds to the mean flow in the Reynolds decomposition and the covariance matrix then corresponds to the Reynolds stress tensor. We will apply the SEnKF method mainly for these analogies with the Reynolds decomposition of turbulent flows.

The used forward model comprises a high-order finite-difference flow solver for the Navier–Stokes equations for the Direct Numerical Simulation (DNS) of turbulent incompressible flow (31) which is thoroughly validated and has been used for several complex flow configurations (32–34) and recently for the study of fluid–structure interaction problems (35, 36). The numerical forecast provided by the forward model are decomposed in an expectation value (ensemble-average) and its fluctuations. Such expectation values can be interpreted as a RANS flow field and the associated covariance matrix as the Reynolds stresses (RSS). In that spirit, our SEnKF algorithm applies a correction to the numerical forecast based on its covariance (RSS) and on a set of observed data and their associated covariances.

The remainder of this paper is organized as follows: the section 2 provides a brief description of the theoretical background of turbulent flows, EnKF and SEnKF approaches, and the proposed DA methodology that will estimate the enhanced flow fields. In the section 3, the proposed DA methodology is applied to three cases: Unsteady flow past a circular cylinder confined in a channel; Wall-bounded turbulent flow in a channel; Flow downstream of an aortic valve. The paper concludes with a discussion on the accuracy, efficiency, and versatility of the DA methodology based on these results.

2. METHODS

2.1. Ensemble Averaging

We model incompressible flow of a homogeneous Newtonian fluid with the Navier–Stokes equations

$$\vec{\nabla} \cdot \vec{u}(\vec{x}, t) = 0, \quad (1)$$

$$\frac{\partial \vec{u}(\vec{x}, t)}{\partial t} + \left(\vec{u}(\vec{x}, t) \cdot \vec{\nabla} \right) \vec{u}(\vec{x}, t) + \frac{1}{\rho} \vec{\nabla} p(\vec{x}, t) - \nu \nabla^2 \vec{u}(\vec{x}, t) - \vec{f}(\vec{x}, t) = 0, \quad (2)$$

where $\vec{x} = \{x, y, z\}$ are Cartesian coordinates, and t is the time. The variables ρ and ν are the fluid density and the kinematic viscosity, respectively; p is the fluid pressure; \vec{u} and \vec{f} are the fluid velocity and the volumetric forcing, respectively.

In the case of turbulent flows, the pressure and velocity fields present chaotic, unsteady changes due to the nonlinear nature of the system and exhibit strong sensitivity to the initial conditions. In practice, this leads to non-reproducible flow fields despite the deterministic nature of the Navier–Stokes (Equation 2) and a statistical approach is often used to study turbulent flow

systems. Therefore, we will not aim at predicting exact turbulent fluctuations from noisy observations. Rather, we will formulate a data assimilation scheme to predict statistical properties of the turbulent flow comprising the mean flow and the second moment of the turbulent fluctuations.

The expectation value of a specific turbulent flow quantity may be estimated from an ensemble of multiple realizations of the flow field, i.e., multiple states-of-system. Each realization r of the flow field $\vec{u}(\vec{x}, t)_{(r)}$ can be decomposed in an ensemble-averaged field $\langle \vec{u}(\vec{x}, t) \rangle$ and the fluctuations $\vec{u}'(\vec{x}, t)_{(r)}$ according to the Reynolds decomposition

$$\vec{u}(\vec{x}, t)_{(r)} = \langle \vec{u}(\vec{x}, t) \rangle + \vec{u}'(\vec{x}, t)_{(r)}, \quad (3)$$

with the ensemble-average on the s states-of-system defined as

$$\langle \vec{u}(\vec{x}, t) \rangle := \lim_{s \rightarrow \infty} \frac{1}{s} \sum_{r=1}^s \vec{u}(\vec{x}, t)_{(r)}. \quad (4)$$

This leads to the following Reynolds-averaged Navier–Stokes (RANS) equations for incompressible flows:

$$\vec{\nabla} \cdot \langle \vec{u}(\vec{x}, t) \rangle = 0 \quad (5)$$

$$\begin{aligned} \frac{\partial \langle \vec{u}(\vec{x}, t) \rangle}{\partial t} + \left(\langle \vec{u}(\vec{x}, t) \rangle \cdot \vec{\nabla} \right) \langle \vec{u}(\vec{x}, t) \rangle + \frac{1}{\rho} \vec{\nabla} \langle p(\vec{x}, t) \rangle \\ - \nu \nabla^2 \langle \vec{u}(\vec{x}, t) \rangle - \langle \vec{f}(\vec{x}, t) \rangle \\ = -\vec{\nabla} \cdot \langle \vec{u}'(\vec{x}, t) \vec{u}'(\vec{x}, t)^T \rangle. \end{aligned} \quad (6)$$

Equations (5) and (6) have the same form as the original (Equations 1, 2) except for the additional term on the right-hand side of Equation (6) comprising the so-called Reynolds Stresses (RSS) $-\rho \langle \vec{u}'(\vec{x}, t) \vec{u}'(\vec{x}, t)^T \rangle$.

In pulsatile flows, the flow configurations are repeated with period T , and a generic time t_n can be written as

$$t_n = t_{(r),\phi} = t_\phi + (r - 1) T, \quad (7)$$

where t_ϕ is the phase time and r is the number of the pulse. The description of the pulsatile dynamics can be reduced to the study of the flow configurations of the basic pulse period. Nonetheless, in presence of turbulence, the periodically subsequent configurations of the pulsatile flow will show differences due to the turbulent fluctuations. Information about the turbulent dynamics of pulsatile flows can be captured decomposing the flow field by phase-averaging, that is obtained by building the ensemble of states-of-system from the periodically subsequent configurations. Thus, the ensemble definition (Equation 4) assumes the following form

$$\langle \vec{u}(\vec{x}, t_n) \rangle_\phi := \lim_{s \rightarrow \infty} \frac{1}{s} \sum_{r=1}^s \vec{u}(\vec{x}, t_\phi + (r - 1) T) \quad (8)$$

where $\langle \cdot \rangle_\phi$ is the phase-average operation on the periodically subsequent configurations related to the basic time t_ϕ . The

pulsatile turbulent dynamics is then described at each basic time t_ϕ by the phase-averaged flow field $\langle \bar{u}(\bar{x}, t_\phi) \rangle_\phi$ and its covariance $\langle \bar{u}'(\bar{x}, t_\phi) \bar{u}'(\bar{x}, t_\phi)^T \rangle_\phi$.

In turbulent configurations presenting a statistically-steady behavior, a time-average corresponds to an ensemble-average. Thus, the ensemble definition (Equation 4) assumes the following form

$$\langle \bar{u}(\bar{x}, t) \rangle_t := \lim_{s \rightarrow \infty} \frac{1}{s} \sum_{r=1}^s \bar{u}(\bar{x}, t_0 + (r-1)\Delta\tau) \quad (9)$$

where $\langle \cdot \rangle_t$ is the time-average operation, t_0 is the initial time value and $\Delta\tau$ a chosen time increment. The statistically-steady turbulent dynamics is then described by the time-averaged flow field $\langle \bar{u}(\bar{x}, t) \rangle_t$ and its covariance $\langle \bar{u}'(\bar{x}, t) \bar{u}'(\bar{x}, t)^T \rangle_t$. Note that the definition of the time-average operation (Equation 9) is formally identical to the definition ensemble average (Equation 8) if the number of phases describing the flow is equal to 1 and $\Delta\tau$ corresponds the time period T .

In spatially-homogeneous turbulent flows the turbulent features are statistically the same in the flow domain such that the ensemble-average (Equation 4) can be replaced by a volume-average,

$$\overline{\bar{u}(\bar{x}, t)} := \frac{1}{V} \iiint_V \bar{u}(\bar{\xi}, t) d\bar{\xi} \quad (10)$$

where $\overline{\cdot}$ stands for the volume-average operation over the volume V . The spatially-homogeneous turbulent dynamics is then described by the volume-averaged flow field $\overline{\bar{u}(\bar{x}, t)}$ and its associated covariance $\overline{\bar{u}'(\bar{x}, t) \bar{u}'(\bar{x}, t)^T}$.

In summary, the ensemble used for statistical characterization of a turbulent flow can be built in different ways depending on the physical behavior of the turbulent flow configuration (statistically-steady, periodic, spatially-homogeneous) leading to different forms of Equation (4). Equations (8)–(10) can be seen as special cases of the following most general form of the ensemble definition Equation (4)

$$\langle \overline{\bar{u}(\bar{x}, t_n)} \rangle_\phi := \lim_{s \rightarrow \infty} \frac{1}{s} \sum_{r=1}^s \overline{\bar{u}(\bar{x}, t_{(r),\phi})}. \quad (11)$$

2.2. Stochastic Ensemble Kalman Filtering Approach

The SENKF approach is used to estimate the state of the system by filtering an ensemble forecast with observations over time. Hereafter, a brief introduction to KF algorithms and then to the SENKF is presented. Details of the algorithms are available in the original papers (2, 4, 5).

The true state-of-system at time t_n is denoted by the state-vector \bar{u}_n . The observation of this state is denoted by the vector \bar{d}_n which depends on \bar{u}_n through the observation operator \mathbf{H} that describes the measurement tool used for the acquisition of the data, such that

$$\bar{d}_n = \mathbf{H}\bar{u}_n + \bar{r}_n, \quad (12)$$

where \bar{r}_n is the observation noise described by a normal distribution $\bar{r}_n \sim \mathcal{N}(\bar{0}, \mathbf{R}_n)$ with zero mean and covariance \mathbf{R}_n . The forward model is defined through the operator \mathbf{M}_n such that

$$\bar{u}_n = \mathbf{M}_n \bar{u}_{n-1} + \bar{q}_n, \quad (13)$$

where the error \bar{q}_n follows a normal distribution $\bar{q}_n \sim \mathcal{N}(\bar{0}, \mathbf{Q}_n)$ with zero mean and covariance \mathbf{Q}_n . The operator \mathbf{M}_n models the physical behavior of the system and can be of linear or nonlinear nature.

To estimate the state at time level n , we assume that the state of the system \bar{u}_{n-1} given all past observations to that time step $\bar{d}_{1:n-1}$ follows a normal distribution with mean $\bar{\mu}_{n-1}^a$ and with covariance \mathbf{P}_{n-1}^a , such that

$$\bar{u}_{n-1} | \bar{d}_{1:n-1} \sim \mathcal{N}(\bar{\mu}_{n-1}^a, \mathbf{P}_{n-1}^a). \quad (14)$$

Starting from this assumption on the previous time level $n-1$, we can estimate a forecast of the state vector \bar{u}_n as

$$\bar{u}_n | \bar{d}_{1:n-1} \sim \mathcal{N}(\bar{\mu}_n^f, \mathbf{P}_n^f), \quad (15)$$

where the superscript f stands for the *forecast* step that is the prior update obtained from the forward model. In the following, the superscript a denotes the *analysis* step that is the posteriori update of the system obtained by taking into account also the observations at that time step.

The basic idea of the EnKF algorithms (4) is to estimate $\bar{\mu}_n^f$ and \mathbf{P}_n^f from an ensemble of s states $\bar{u}_{(r),n}^f$ with $r = 1, 2, \dots, s$,

$$\bar{u}_{(r),n}^f = \mathbf{M}(\bar{u}_{(r),n-1}^a, t_n) \quad (16)$$

$$\bar{\mu}_n^f = \frac{1}{s} \sum_{r=1}^s \bar{u}_{(r),n}^f \quad (17)$$

$$\mathbf{P}_n^f = \frac{1}{s} \sum_{r=1}^s (\bar{u}_{(r),n}^f - \bar{\mu}_n^f) (\bar{u}_{(r),n}^f - \bar{\mu}_n^f)^T. \quad (18)$$

If the observations \bar{d}_n can be treated as the result of an ensemble of s samples $\bar{d}_{(r),n}$, the observation noise covariance is replaced by a sample covariance, which leads to the following relations for the SENKF algorithm (5):

$$\bar{d}_{(r),n} = \mathbf{H}\bar{u}_{(r),n} + \bar{r}_{(r),n} \quad (19)$$

$$\bar{d}_n = \frac{1}{s} \sum_{r=1}^s \bar{d}_{(r),n} \quad (20)$$

$$\mathbf{R}_n = \frac{1}{s} \sum_{r=1}^s (\bar{d}_{(r),n} - \bar{d}_n) (\bar{d}_{(r),n} - \bar{d}_n)^T. \quad (21)$$

After the prior update obtained from the forecast step (Equations 16–18), the SENKF algorithms proceeds with the analysis step

which updates the state-vector using the observations:

$$\vec{u}_{(r),n}^a = \vec{u}_{(r),n}^f + \mathbf{K}_n \left(\vec{d}_{(r),n} - \mathbf{H} \vec{u}_{(r),n}^f \right) \quad (22)$$

$$\vec{\mu}_n^a = \frac{1}{s} \sum_{r=1}^s \vec{u}_{(r),n}^a \quad (23)$$

$$\mathbf{P}_n^a = \frac{1}{s} \sum_{r=1}^s \left(\vec{u}_{(r),n}^a - \vec{\mu}_n^a \right) \left(\vec{u}_{(r),n}^a - \vec{\mu}_n^a \right)^T. \quad (24)$$

Equation (22) defines the correction of the forecast $\vec{u}_{(r),n}^f$ to the state estimate $\vec{u}_{(r),n}^a$ through the so-called Kalman gain \mathbf{K}_n which depends on the observation uncertainty \mathbf{R}_n and the uncertainty of the forecast estimate \mathbf{P}_n^f according to

$$\mathbf{K}_n = \mathbf{P}_n^f \mathbf{H}^T \left(\mathbf{H} \mathbf{P}_n^f \mathbf{H}^T + \mathbf{R}_n \right)^{-1}. \quad (25)$$

For the limiting case, where the observations are assumed to be perfect, i.e., $\mathbf{H}_n = \mathbf{I}$ and $\mathbf{R}_n \rightarrow \mathbf{0}$ (where \mathbf{I} is the identity matrix), the Kalman gain reduces to $\mathbf{K}_n = \mathbf{I}$ such that the analysis step becomes $\vec{u}_{(r),n}^a = \vec{d}_{(r),n}$. This limiting case illustrates that the state estimate $\vec{u}_{(r),n}^a$ tends to be dominated by the observations $\vec{d}_{(r),n}$ if \mathbf{R}_n is small. In contrast, $\mathbf{K}_n \rightarrow \mathbf{0}$ if the prior covariance \mathbf{P}_n^f tends to zero, i.e., if the level of forecast certainty is high. In that case, the filter ignores the observations and $\vec{u}_{(r),n}^a \approx \vec{u}_{(r),n}^f$.

2.3. Data Assimilation Methodology for Pulsatile Turbulent Flows

In the following, we extend the SEnKF approach to formulate a DA methodology for pulsatile, turbulent flows.

2.3.1. Data Acquisition

In the present context, we assume that observations of pulsatile, turbulent flows are available as voxel-based data with voxel size h . Often small turbulent length scales cannot be captured by such observations. The measured data can be seen as the result of a volumetric average over the voxel volume according to Equation (10). Moreover, if the observations are run for p repetitive pulses, the voxel data can also be phase averaged to obtain estimates of mean and a covariance according to Equation (11). Each observed voxel (v)-based data $\vec{d}_{v,(s),\phi}$ measured in the pulse s has an error $\vec{r}_{v,(s),\phi}$ whose covariance will be considered as the result of a voxel-phase-average. This leads to the following definitions of the data $\vec{d}_{v,(s),\phi}$, the associated mean $\vec{\mu}_{v,\phi}$ and sample covariance $\mathbf{r}_{v,\phi}$:

$$\vec{d}_{v,(s),\phi} = \overline{\vec{u}(\vec{x}, t_{(s),\phi})} \quad (26)$$

$$\vec{\mu}_{v,\phi} = \overline{\langle \vec{u}(\vec{x}, t_\phi) \rangle_\phi} = \frac{1}{p} \sum_{s=1}^p \vec{d}_{v,(s),\phi}, \quad (27)$$

$$\begin{aligned} \mathbf{r}_{v,\phi} &= \overline{\langle \vec{u}(\vec{x}, t_\phi) \rangle_\phi}' \overline{\langle \vec{u}(\vec{x}, t_\phi) \rangle_\phi}^T \\ &= \frac{1}{p} \sum_{s=1}^p \left(\vec{d}_{v,(s),\phi} - \vec{\mu}_{v,\phi} \right) \left(\vec{d}_{v,(s),\phi} - \vec{\mu}_{v,\phi} \right)^T. \end{aligned} \quad (28)$$

In the classical relation (Equation 21), the covariance matrix is written with respect to the entire state-vector \vec{d}_n . Here for simplicity we assume that covariance between velocities of different voxels is zero; thus, we define the covariance matrix \mathbf{r}_v independently for each voxel v .

2.3.2. Forecast Solution

The forecast solution is computed using a Navier–Stokes solver for direct numerical simulation (DNS). In the EnKF formalism, this solver represents the nonlinear forward model ($\mathbf{M} = \text{DNS}$) between two subsequent update steps $n - 1$ and n such that

$$\vec{u}_n = \text{DNS}(\vec{u}_{n-1}, t_n) + \vec{q}_n, \quad (29)$$

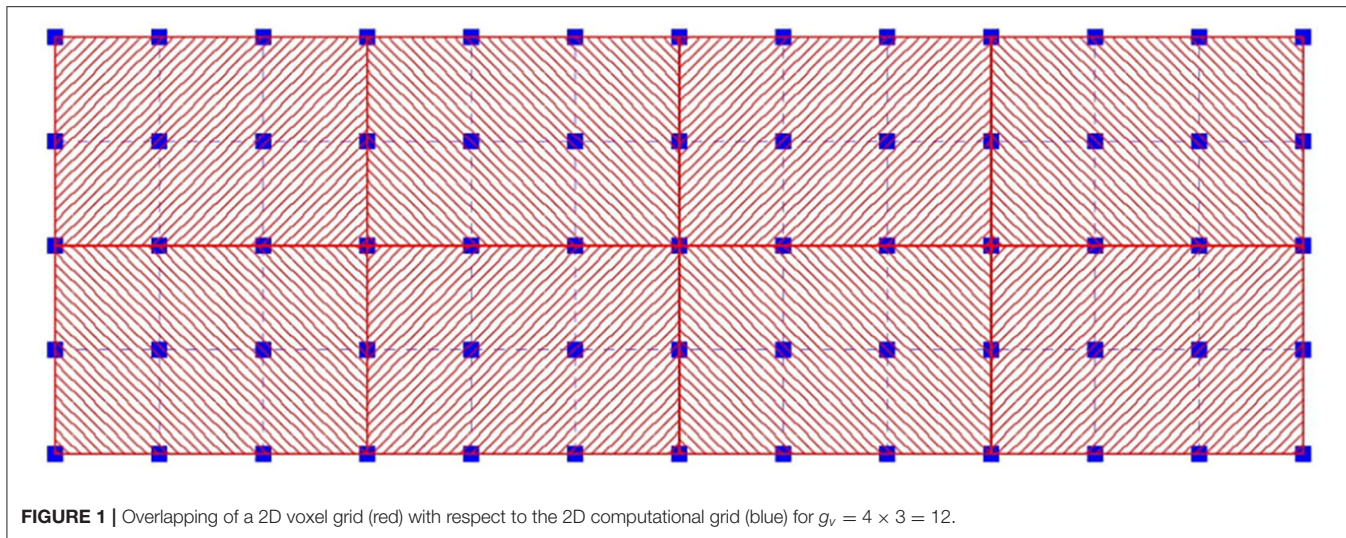
$$\vec{u}_{(r),n}^f = \text{DNS}(\vec{u}_{(r),n-1}^a, t_n). \quad (30)$$

The Kalman time step $\Delta t_{KF} = t_{n+1} - t_n$ between two subsequent updates should not be confused with the DNS time step of the Navier–Stokes solver. Even though the DNS Equation 30 returns a deterministic solution $\vec{u}_{(r),n}^f$ affected only by a numerical error, we know that turbulent flows have a stochastic behavior that can be investigated *a posteriori* by using an ensemble-average approach in order to evaluate the effect of the turbulent fluctuations. The deterministic solution $\vec{u}_{(r),n}^f$ differs from the ground-truth solution \vec{u}_n of Equation 29 because of the error \vec{q}_n which is composed of three contributions of different nature: a numerical error \vec{q}_{nNUM} due to the limited numerical accuracy of the discretized Navier–Stokes equations; an uncertainty \vec{q}_{nTUR} due to the stochastic nature of turbulent fluctuations; and a modeling error \vec{q}_{nMOD} due to the uncertainty error (e.g., unknown boundary conditions, viscosity) typical of blood flow physical modeling:

$$\vec{q}_n = \vec{q}_{nNUM} + \vec{q}_{nTUR} + \vec{q}_{nMOD}. \quad (31)$$

Here, the numerical error \vec{q}_{nNUM} is assumed to be negligible, because in a DNS its magnitude must be lower than the magnitude of turbulent fluctuations. In the case of blood flow simulations, the error \vec{q}_{nMOD} will add additional uncertainty that will affect the forecast solution.

In the present work, the DNS data $\vec{u}_{(r),n}^f$ is considered as a stochastic realization of the ground truth, i.e., it comprises a mean flow component and a random turbulent fluctuation. In the language of the Kalman filter theory, the mean flow is the ground truth and the Reynolds stress tensor is the covariance matrix \mathbf{P} of the forecast noise. This interpretation neglects the effect of numerical noise and modeling uncertainty. The appropriate choice of the forecast ensemble for the SEnKF algorithm is important for obtaining a good filter for the specific flow configuration. For pulsatile turbulent configurations, we propose to build the ensemble of states from multiple pulses, i.e., each pulse is considered a different realization. First, the pulse period T is divided in a number of intervals N_ϕ defined by the basic time $t_\phi = t_1, t_2, \dots, t_{N_\phi}$ with $t_{N_\phi} = T$. We generate an increasingly large ensemble by running the forward model for multiple pulses. This is a significant deviation from the standard SEnKF concept, where the ensemble is typically built from multiple computations of the forward model with slightly different initial conditions.



In accordance with Equations (16)–(18), the forecast state-vector, its mean and associated sample covariance at each computational grid node g are then computed as:

$$\vec{u}_{g;(s),\phi}^f \equiv \vec{u}^f(\vec{x}_g, t_{(s),\phi}), \quad (32)$$

$$\vec{\mu}_{g;(s),\phi}^f \equiv \langle \vec{u}^f(\vec{x}_g, t) \rangle_{(s),\phi} = \frac{1}{s} \sum_{r=1}^s \vec{u}_{g;(r),\phi}^f, \quad (33)$$

$$\langle \vec{u}^f(\vec{x}_g, t) \vec{u}^f(\vec{x}_g, t) \rangle_{(s),\phi} = \frac{1}{s} \sum_{r=1}^s \left(\vec{u}_{g;(r),\phi}^f - \vec{\mu}_{g;(s),\phi}^f \right) \left(\vec{u}_{g;(r),\phi}^f - \vec{\mu}_{g;(s),\phi}^f \right)^T. \quad (34)$$

Here for simplicity we assume that covariance between velocities of different grid nodes is zero. Thus, we define the covariance matrix \mathbf{p}_g^f independently for each grid node.

The ensemble will be very small in the beginning and only becomes a statistically useful sample of realizations after several pulses. In the first pulse of the numerical simulation ($s = 1$), the sample covariance $\mathbf{p}_{g;(1),\phi}^f$ would be equal to zero according to Equation (34). Therefore, the covariance $\mathbf{p}_{g;(s),\phi}^f$ is computed in the following way:

$$\mathbf{p}_{g;(s),\phi}^f = \frac{1}{s} \left(D + \sum_{r=1}^s \left(\vec{u}_{g;(r),\phi}^f - \vec{\mu}_{g;(s),\phi}^f \right) \left(\vec{u}_{g;(r),\phi}^f - \vec{\mu}_{g;(s),\phi}^f \right)^T \right), \quad (35)$$

where D is a diagonal matrix whose elements are on order of magnitude larger than the expected covariance one of the specific flow configuration. D ensures the positiveness of $\mathbf{p}_{g;(s),\phi}^f$ at the beginning of the simulation, and then its influence disappears for increasingly larger ensembles.

2.3.3. Data Assimilation Algorithm

The proposed methodology aims at using observed data on a voxel-grid in combination with the forecast solution computed by DNS on a finer computational grid. To this end, the voxels m_v of the voxel grid are overlapped with the computational grid nodes m_g , where the state-vector at each time t_n is defined on each grid node g as $\vec{u}_{g;n} = \{u_x, u_y, u_z\}_{g;n}$. For Cartesian grids, g_v computational grid-nodes will lie inside each voxel v ; and it is possible that a grid node g belongs to more than one voxel, i.e., the grid node is located exactly on the borders of adjacent v_g voxels (Figure 1).

In the present context, $\vec{d}_{v;(s),\phi}$ is the value that the solution would have on all computational grid nodes g_v lying inside the voxel v , if the data $\vec{d}_{v;(s),\phi}$ was constant over the whole voxel volume. It is clear that this is generally not the case. The discrete form of the observation operator \mathbf{H} that maps the voxel-based data to the grid is a $3m_v \times 3m_g$ matrix assuming the following form:

$$[H_{ij}]_{v,g} = \phi_{v,g} \delta_{ij} \quad (36)$$

where δ_{ij} is the Kronecker delta, $i, j = 1, 2, 3$ and

$$\phi_{v,g} = \begin{cases} \frac{1}{g_v}, & \text{if } g \text{ lies inside } v, \\ 0, & \text{otherwise.} \end{cases} \quad (37)$$

Solving the analysis step to update the entire state-vector $\vec{u}_{1:m_g;(s),\phi}$ requires the computation on huge matrices \mathbf{P} , \mathbf{R} and \mathbf{H} of size $3m_g \times 3m_g$, $3m_v \times 3m_v$, and $3m_v \times 3m_g$, respectively. However, we can exploit the block-structure of \mathbf{H} comprising of blocks \mathbf{h} of size $3 \times 3g_v$ which correspond to single voxels. From an algorithmic point of view, this allows to only use m_v dense matrices \mathbf{h} instead of the huge sparse matrix \mathbf{H} which dramatically reduces the

computational cost. There, the Kalman gain term is written as

$$\vec{k}_{g;(s),\phi} = \sum_{v=1}^{m_v} \Psi_{g,v} \left[\mathbf{p}_{1:g_v;(s),\phi}^f \mathbf{h}^T \left(\mathbf{h} \mathbf{p}_{1:g_v;(s),\phi}^f \mathbf{h}^T + \mathbf{r}_{v,\phi} \right)^{-1} \left(\vec{d}_{v;(s),\phi} - \mathbf{h} \vec{u}_{1:g_v;(s),\phi}^f \right) \right]_g, \tag{38}$$

where $\Psi_{g,1:m_v}$ are weights for interpolating data between the discrete values of the Kalman gain obtained from the different m_v voxels to the grid node g . Since each grid-node solution depends only on the data of the v_g voxels, we define

$$\Psi_{g,v} = \begin{cases} \frac{1}{v_g}, & \text{if } g \text{ lies inside } v, \\ 0, & \text{otherwise.} \end{cases} \tag{39}$$

Finally, the updated solution $\vec{u}_{(s),\phi}^a$ is computed in the analysis step of the SEnKF as

$$\vec{u}_{g;(s),\phi}^a = \vec{u}_{g;(s),\phi}^f + \vec{k}_{g;(s),\phi}. \tag{40}$$

In conclusion, the proposed methodology deals with data obtained by observations on coarse grids, and enhances them by means of the SEnKF algorithm, obtaining a new set of data on finer grids. The entire DA methodology is summarized in the Algorithm 1.

2.4. Direct Numerical Simulation

All numerical simulations have been performed using a high-order Navier–Stokes solver for the DNS of incompressible flows. The governing Equations (1) and (2) are discretized on a staggered Cartesian grid using sixth-order finite-difference schemes. A multigrid method is used for solving the Poisson problem and time integration is performed using a three-step Runge–Kutta scheme. Details of the implementation are given in Henniger et al. (31).

The simulations use a combination of periodic boundary conditions and the fringe forcing technique (37) which damps incoming flow disturbances and forces the flow field toward a desired velocity profile by applying a suitable volume force \vec{f} of the right-hand side of the Navier–Stokes (Equation 1). The general form of the fringe forcing is given by

$$\vec{f}(\vec{x}, t) = \lambda(\vec{x}) \left(\vec{U}(\vec{x}, t) - \vec{u}(\vec{x}, t) \right) \tag{41}$$

where $\lambda(\vec{x})$ is the fringe function and $\vec{U}(\vec{x}, t)$ is the desired velocity to be imposed. The fringe function is non-zero only within the so-called fringe region which is typically located at the edge of the computational domain. In the physically relevant regions of the DNS, the fringe function is zero.

This technique has been exploited to impose inflow velocity profiles, to model outflow conditions and to model complex geometries by setting $\vec{U} = 0$ within immersed objects. The fringe forcing has similarities to the Kalman

Algorithm 1: Data Assimilation algorithm for computing enhanced voxel-based data

```

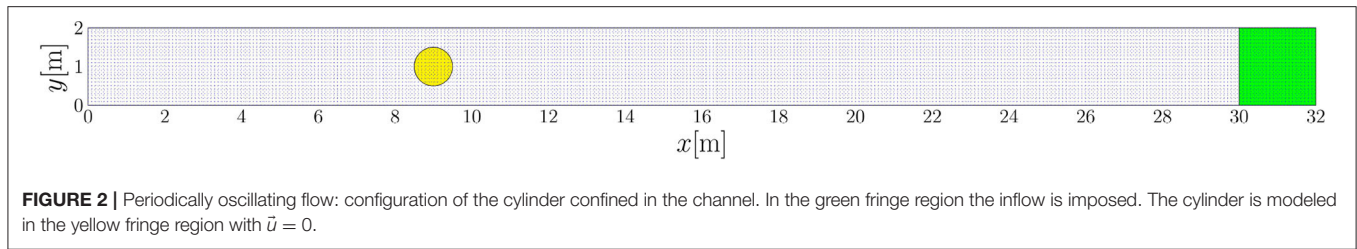
- Initialization:  $\vec{u}_{g;(0),N_\phi}^a = \vec{u}_{i.c.}$  (initial conditions)
- Data Assimilation:
for  $s = 1$  to  $p$  do
  for  $\phi = 1$  to  $N_\phi$  do
     $\vec{u}_{g;(s),0}^a = \vec{u}_{g;(s-1),N_\phi}^a$ 
    for  $\phi = 1$  to  $N_\phi$  do
       $\vec{u}_{g;(s),\phi}^f = \text{DNS}(\vec{u}_{(s),\phi-1}^a)$ 
       $\vec{\mu}_{g;(s),\phi}^f = \frac{1}{s} \sum_{r=1}^s \vec{u}_{g;(r),\phi}^f$ 
       $\mathbf{p}_{g;(s),\phi}^f =$ 
       $\frac{1}{s} \left( D + \sum_{r=1}^s \left( \vec{u}_{g;(r),\phi}^f - \vec{\mu}_{g;(s),\phi}^f \right) \left( \vec{u}_{g;(r),\phi}^f - \vec{\mu}_{g;(s),\phi}^f \right)^T \right)$ 
       $\vec{k}_{g;(s),\phi} =$ 
       $\sum_{v=1}^{m_v} \Psi_{g,v} \left[ \mathbf{p}_{1:g_v;(s),\phi}^f \mathbf{h}^T \left( \mathbf{h} \mathbf{p}_{1:g_v;(s),\phi}^f \mathbf{h}^T + \mathbf{r}_{v,\phi} \right)^{-1} \right.$ 
       $\left. \left( \vec{d}_{v;(s),\phi} - \mathbf{h} \vec{u}_{1:g_v;(s),\phi}^f \right) \right]_g$ 
       $\vec{u}_{g;(s),\phi}^a = \vec{u}_{g;(s),\phi}^f + \vec{k}_{g;(s),\phi}$ 
    for  $\phi = 1$  to  $N_\phi$  do
      Result:  $\vec{\mu}_{g;\phi}^a = \vec{u}^a(\vec{x}_g, t) = \frac{1}{p} \sum_{r=1}^p \vec{u}_{g;(r),\phi}^a$ 
      Result:  $\mathbf{p}_{g;\phi}^a = \langle \vec{u}^a(\vec{x}_g, t) \vec{u}^a(\vec{x}_g, t)^T \rangle_{g;\phi} =$ 
       $\frac{1}{p} \sum_{r=1}^p \left( \vec{u}_{g;(r),\phi}^a - \vec{u}_{g;\phi}^a \right) \left( \vec{u}_{g;(r),\phi}^a - \vec{u}_{g;\phi}^a \right)^T$ 

```

analysis step (Equation 22), where the Kalman gain K acts like a fringe function and the observation \vec{d} corresponds to the desired velocity \vec{U} . However, the fringe forcing is added to the right-hand side of a differential equation and drives the solution toward \vec{u} with a time scale $1/\lambda$, whereas the Kalman update is an algebraic equation which nudges the solution toward the observation at a time scale $\Delta t_{KF}/\|K\|$, where Δt_{KF} is the time between two subsequent analysis steps.

3. RESULTS

In the present study, all observations are available on Cartesian voxel grids and obtained either by experiments or other numerical simulations (synthetic data). In case of synthetic data, the voxel data are extracted from the computational grid, and each voxel comprises multiple g_v grid nodes of the computational grid. The observations have been run for p multiple pulses, and the synthetic voxel data are collected in order to obtain a mean and a covariance estimate in accordance with Equation (11). Each voxel data $\vec{d}_{v;(s),\phi}$ computed in the specific pulse s has an error $\vec{r}_{v;(s),\phi}$ whose covariance will be considered as the result of a voxel-and-phase-average; Equation (26) assumes the



following form:

$$\vec{d}_{v;(s),\phi} \equiv \overline{\vec{u}(\vec{x}, t_{(s),\phi})} = \frac{1}{g_v} \sum_{g=1}^{g_v} \vec{u}_{g;(s),\phi}. \quad (42)$$

The methodology has been validated for test problems of increasing complexity, starting with a periodic and non-turbulent configuration, then extending the application of the methodology to a turbulent configuration, and finally applying the methodology to a pulsatile and turbulent configuration. Results have been evaluated comparing global flow-related parameters and local profiles with respect to the ground truth and observed data.

The first flow related parameter that has been chosen is the friction Reynolds number Re_τ which is defined using the friction velocity U_τ ,

$$Re_\tau = \frac{U_\tau L_{ref}}{\nu} = U_\tau \frac{Re}{U_{ref}} = Re \sqrt{\frac{1}{Re U_{ref}} \left\langle \left\langle \frac{\partial u_x(\vec{x}, t)}{\partial y} \right\rangle \right\rangle}, \quad (43)$$

where Re is the Reynolds number based on reference values (U_{ref} , L_{ref}) of the flow configuration, the average $\langle \langle \cdot \rangle \rangle$ is calculated at the specific time t over the wall. The velocity gradient is discretized with a first-order finite difference scheme on the computational grid. Mean velocity and RSS profiles are commonly described in wall units, i.e., by dimensionless variables y^+ and u^+ obtained by normalization with respect to friction velocity U_τ and flow parameters,

$$u^+ = \frac{u}{U_\tau}, \quad y^+ = \frac{yU_\tau}{\nu}, \quad (44)$$

where y is the distance to the wall and u is the velocity component parallel to the wall.

The turbulent kinetic energy (TKE) is a second parameter used to quantify the turbulence in the bulk flow. We define TKE within a given volume interest V , such that

$$TKE(t) = \frac{1}{2} \overline{u'_x(\vec{x}, t)^2 + u'_y(\vec{x}, t)^2 + u'_z(\vec{x}, t)^2} \quad (45)$$

Finally, voxel-and-phase(time)-averaged velocity fields and/or RSS fields are presented for the different configurations.

3.1. Periodically Oscillating Flow

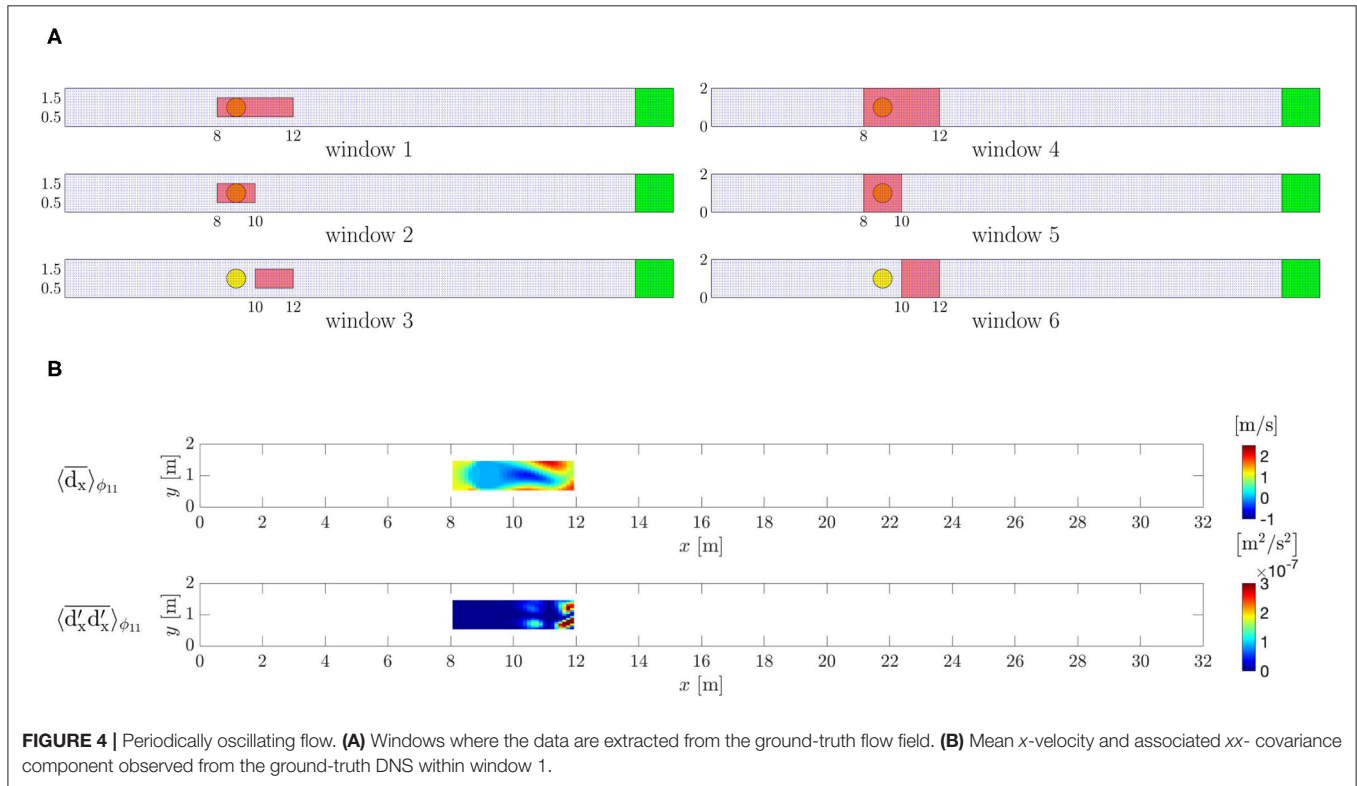
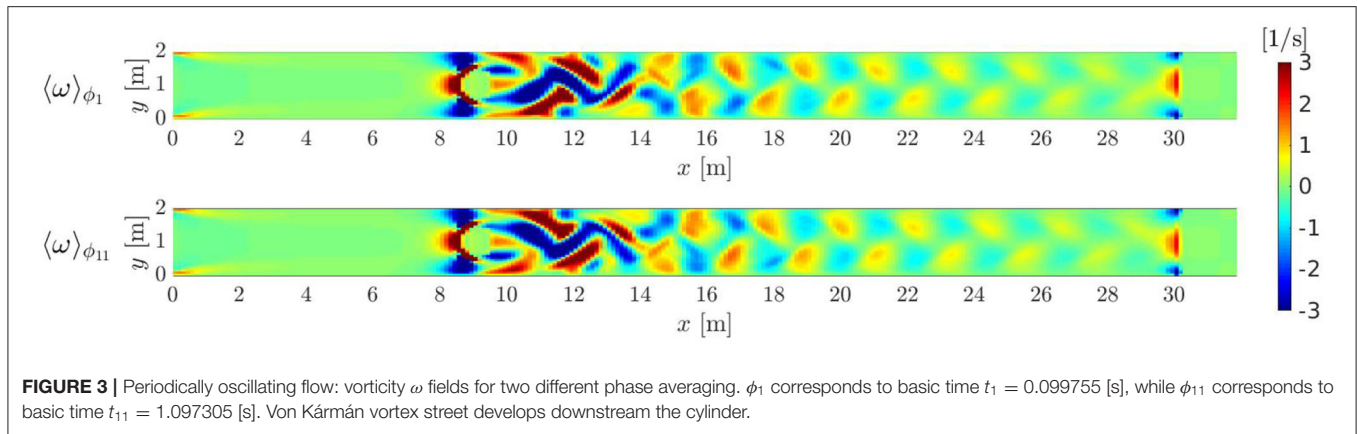
The proposed DA methodology is first validated for an unsteady flow past a cylinder confined in a channel (**Figure 2**). The size of the channel is $32h \times 2h \times h$ where $h = 1$ [m] is the channel half-width. The positive x -direction is the stream-wise direction, the y -axis is perpendicular to the walls at $y = 0$ and $y = 2h$ and the z -axis points in span-wise direction. The diameter of the cylinder is equal to h and its center C has coordinates $(x, y) = (9h, h)$ and an axis parallel to the z -direction. Periodic boundary conditions have been imposed in both the stream-wise and span-wise direction, whereas a no-slip condition is enforced at the top and bottom wall. Fringe forcing at the end of the domain (green region in **Figure 2**) is used to enforce plug flow with velocity $U_x = 1$, $U_y = 0$, $U_z = 0$ [m/s] at the inflow (using $\lambda = 50$ in Equation 41). The presence of the cylinder is modeled by using another fringe region with $\vec{U} = 0$ and $\lambda = 100$ for the grid nodes within the cylinder (yellow region in **Figure 2**). The combination of the plug flow with the no-slip b.c. will produce a thin boundary layer in the green fringe region. The thickness of this boundary layer will increase and the flow will develop a parabolic profile upstream of the cylinder. The Reynolds number is set equal to 150 based on h and the inflow velocity U_x .

3.1.1. Ground Truth

First, a DNS has been run to obtain the ground truth, using a computational grid of $257 \times 33 \times 3$ grid nodes. At the given Reynolds number, the flow past a confined cylinder presents vortex shedding where the vortices detach periodically with a period of $T = 1.9951$ [s] from either side of the body forming a Von Kármán vortex street in the wake of the cylinder (**Figure 3**).

3.1.2. Data Acquisition

Second, the observation data $\vec{d}_{v;(s),\phi}$ have been extracted from the ground-truth DNS. The period T has been divided in $N_\phi = 20$ equal time intervals and the data are collected for $p = 500$ periods starting after $100T$ in order to overcome the initial transient leading to the periodic solution. Six different regions of interest (windows) have been selected to sample the data (**Figure 4A**). Window 1 includes the cylinder and the downstream region from where the vortex shedding originates. Window 2 and window 3 are obtained as the left and right half of window 1, respectively, to study the capability of the proposed method to reconstruct the flow downstream of the cylinder with and without the data of the cylinder itself. The windows 1, 2, and 3 have the lengths 4, 2, and 2, h , respectively. Their height is h and they are centered in the y -direction and extend over the whole z -direction. Extending the windows 1,



2, and 3 to the overall y -domain leads to the windows 4, 5, and 6, respectively. The choice of window 4, 5, and 6 allows to assess whether the availability of data close to the channel walls will improve the flow reconstruction by using the DA algorithm or not.

Each $\vec{d}_{v;(s),\phi}$ is computed for a voxel formed by $g_v = 2 \times 2 \times 2$ grid nodes by averaging the ground-truth solution of the corresponding g_v nodes. These voxel-averaged values are then further phase-averaged in order to obtain voxel-and-phase averaged data $\vec{d}_{v;\phi}$ and the associated covariances $\mathbf{r}_{v;\phi}$ for all the phases, within the specific window (**Figure 4B**). Because the flow is not turbulent, the covariance will be almost zero and its value will just include the covariance of the numerical error of the

DNS solver and the (very low) inaccuracy of the chosen period value T .

3.1.3. Comparison Between Ground Truth and DA Predictions

In the DA predictions only the no-slip b.c. at the walls and the fringe forcing for the inflow are imposed *a priori*. The cylinder fringe region is not used anymore and only the use of the observed data in the SENKF algorithm, will reconstruct the flow field in the wake of the cylinder. The diagonal matrix \mathbf{D} in Equation (35) is set to $\{D_{ii}\} = 1,000$. Six DA predictions, i.e., DA 1, DA 2, ... DA 6, have been run by using the data set extracted

from the six windows of interest, i.e., window 1, window 2, ... window 6, respectively.

The voxel-and-phase averaged velocity profiles $\langle u_x(\vec{x}, t_\phi) \rangle_\phi$ computed after $500T$ are shown in **Figure 5A** (left) along a line parallel to the y -axis and crossing the center of the cylinder at $x = 9$ [m]. When the data window includes the cylinder (windows 1, 2, 4, 5) the SEnKF forces the flow velocities to zero, because the data inside the cylinder have mean and covariance both equal to zero. In contrast, the DA predictions using data from window 3 and 6 show a nearly parabolic velocity profile, which is the velocity profile that the flow assumes due to the inflow and boundary conditions. This indicates that the proposed methodology has limitations in propagating the information from the observed data in upstream direction. **Figure 5A** (right) shows the axial velocity profiles at $x = 15$ [m] in the wake of the cylinder: all the DA predictions present good agreement with the ground truth except for the DA prediction performed by using data from window 2. This shows that the window 2 is too small and do not contain enough data to guide the solution of the DA prediction toward the ground truth. **Figure 5B** shows the xx -covariance component profiles at $x = 15$ [m] (right) and at $x = 9$ [m] (left). DA predictions present an error with respect to the ground truth DNS due to the additional uncertainty related to the cylinder boundary condition and to the convergence history of the DA methodology. On the other hand the information is well propagated in downstream direction by the proposed DA methodology as can be seen in **Figure 5C** for all DA predictions. DA 2 presents a phase-shift with respect to the DNS due to the absence of data in the region surrounding the cylinder which apparently led to a counter-phase vortex detachment.

3.2. Turbulent Channel Flow

Hereafter, the DA methodology is extended to turbulent flows, and validated for turbulent channel flow. To this end, the dimensions of the channel are set to $4\pi h \times 2h \times 2\pi h$ according to Kim et al. (38), see **Figure 6**. The laminar Poiseuille flow $\vec{u} = \{[y(2h - y)/h^2]U_x, 0, 0\}$ having maximum velocity $U_x = 1$ [m/s] has been used as initial condition for the velocity field. The Reynolds number (based on the channel half-width $h = 1$ [m] and the initial maximum velocity U_x) is set to 5,000 which is in the range of the intended final application for blood flow in the aorta (39). Periodic boundary conditions have been imposed in both the stream-wise and span-wise direction, while a no-slip condition is ensured at the top and bottom wall. A constant non-dimensional bulk velocity $U_x^{bk} = 0.667$ [m/s] has been enforced in the stream-wise direction according to Schlatter et al. (40).

3.2.1. Ground Truth

DNS simulations have been run in order to obtain grid-converged solutions that will be considered the ground truth. In order to ensure the transition to turbulence in the channel flow, the initial Poiseuille flow has been perturbed with a two-dimensional (stable) Tollmien–Schlichting (TS) wave with maximum stream-wise velocity amplitude of 3% of the laminar center-line velocity and two superimposed weak oblique (stable) three-dimensional waves with amplitude 0.1% with the same fundamental stream-wise wavelength as the two-dimensional

disturbance. The computation of the TS waves was performed using a standard Chebyshev collocation method involving the solution to the Orr–Sommerfeld and Squire equations (41). Sufficient convergence for Re_τ and TKE was achieved for a resolution of $65 \times 65 \times 65$ points.

3.2.2. Data Acquisition

The data $\vec{d}_{v;t}$ to be used in the DA predictions have been extracted from the flow field of the ground-truth DNS starting after 200 [s] when a statistically-steady turbulent flow had been established. The ensemble of states-of-system is chosen in accordance with Equation (9).

Four different windows have been selected (**Figure 7A**). Window 1 has a size of $(\pi/2)h$ in x -direction and h in y -direction. It is centered in y -direction. Windows 2 and 3 are obtained by extending the window 1 over the whole y - and x -direction, respectively, in order to evaluate the influence of the amount of available data on the flow reconstruction. Window 4, which covers the whole domain has been chosen to mimic the configuration of clinical applications where the data are available for the entire blood vessels. All windows are spread over the whole domain in z -direction.

Data $\vec{d}_{v;t}$ are collected for 4,000 [s] in order to compute the covariance $\mathbf{r}_{v;t}$ in an accurate way onto the voxel-grid. Each $\vec{d}_{v;t}$ is obtained by averaging the ground-truth solution over voxels formed by $g_v = 3 \times 3 \times 3$ adjacent grid-nodes; these voxel-averaged values are then further time-averaged in order to obtain time-and-space averaged data. Because the flow is turbulent, the covariance will be not zero and its value will include both the covariance of the numerical error of the DNS solver and the turbulent fluctuations covariance (**Figure 7B**).

3.2.3. Comparison Between Ground Truth and Data Assimilation Predictions

In the DA predictions the contribution of each voxel-based data is spread onto the g_v grid-nodes according to Equation (38). The initial conditions of the velocity are set to Poiseuille flow without any perturbations such that the flow would not show a transition to turbulence without any additional external forcing due to the SEnKF. The diagonal matrix \mathbf{D} in Equation (35) is set to $\{D_{ii}\} = 10$.

Four DA predictions, i.e., DA 1, DA 2, DA 3, DA 4, have been run by using the data set extracted from the four windows of interest, i.e., window 1, window 2, window 3, window 4, respectively.

The results of the DA predictions have been evaluated by comparing the evolution in time of the friction Reynolds number $\langle Re_\tau \rangle$ and $TKE(t)$ with the ground truth (**Figure 8**). At the beginning of the DA predictions, filtering the initial Poiseuille flow by the observed data leads to a transient evolution of the flow to the statistically-steady turbulent configuration. This transient is different for each window with regard to the mean value to which the solution tends and the time required to overcome the transient. After this transient, Re_τ present the same qualitative behavior in time of the ground truth, even though with different mean value for the different selected data

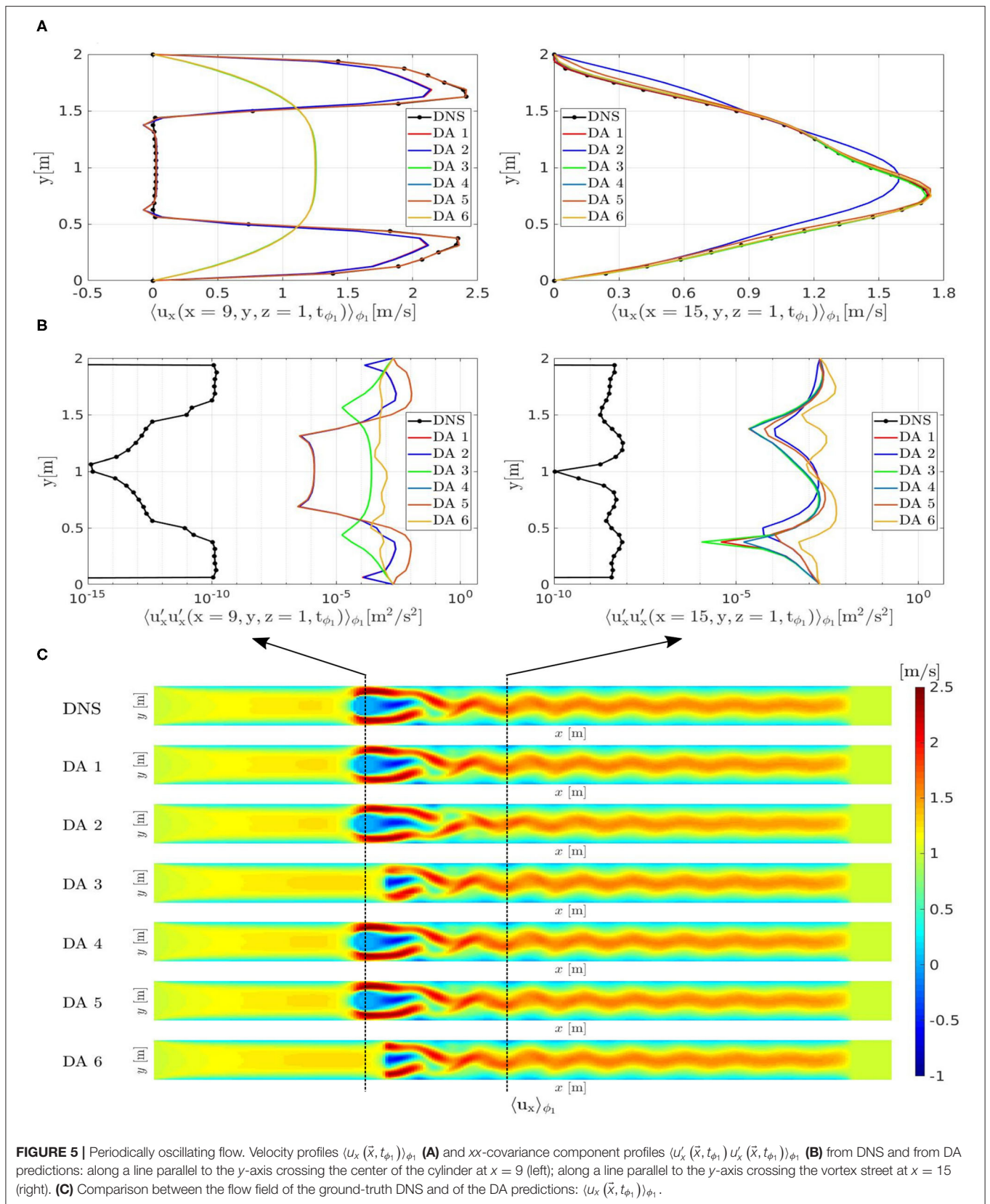
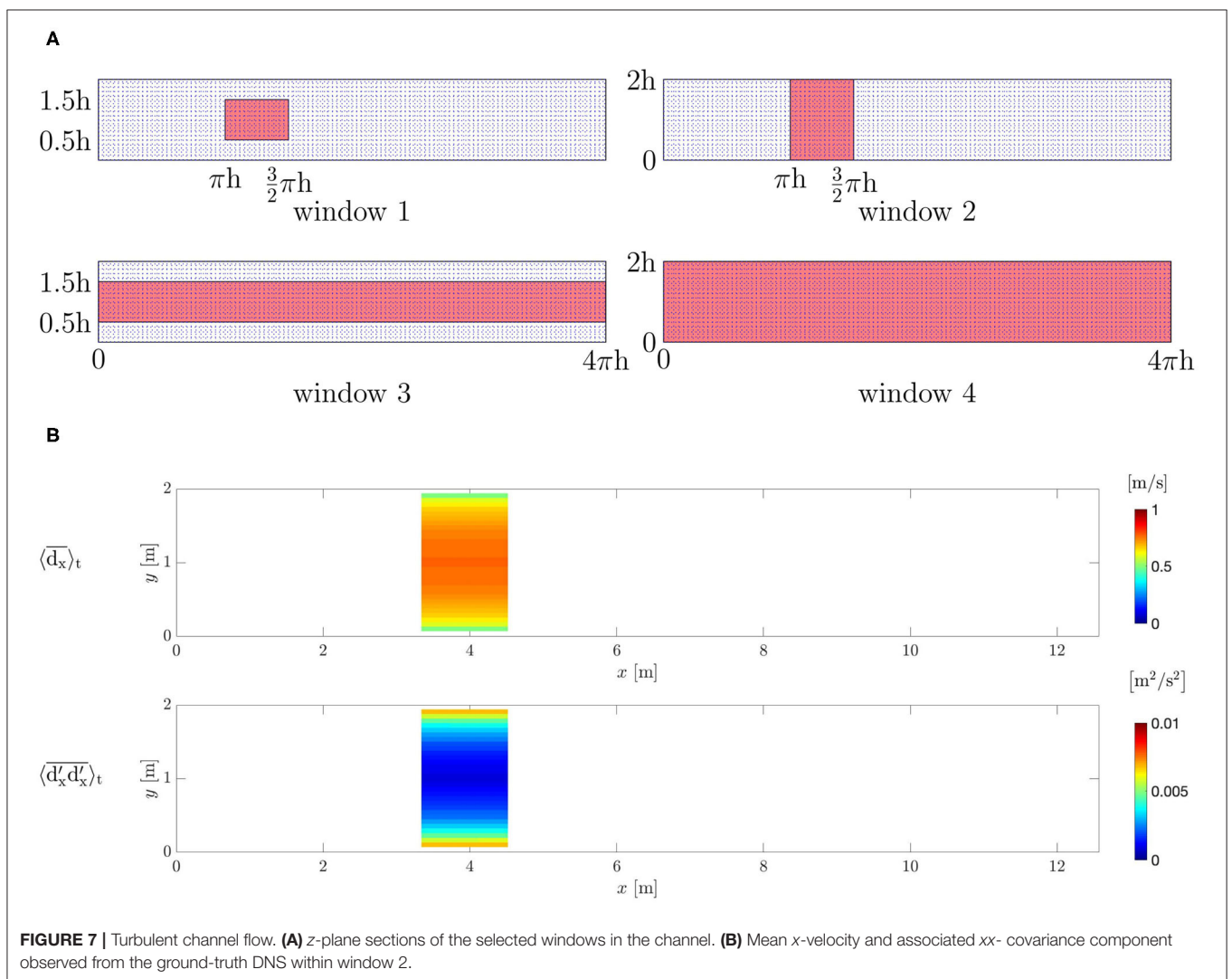
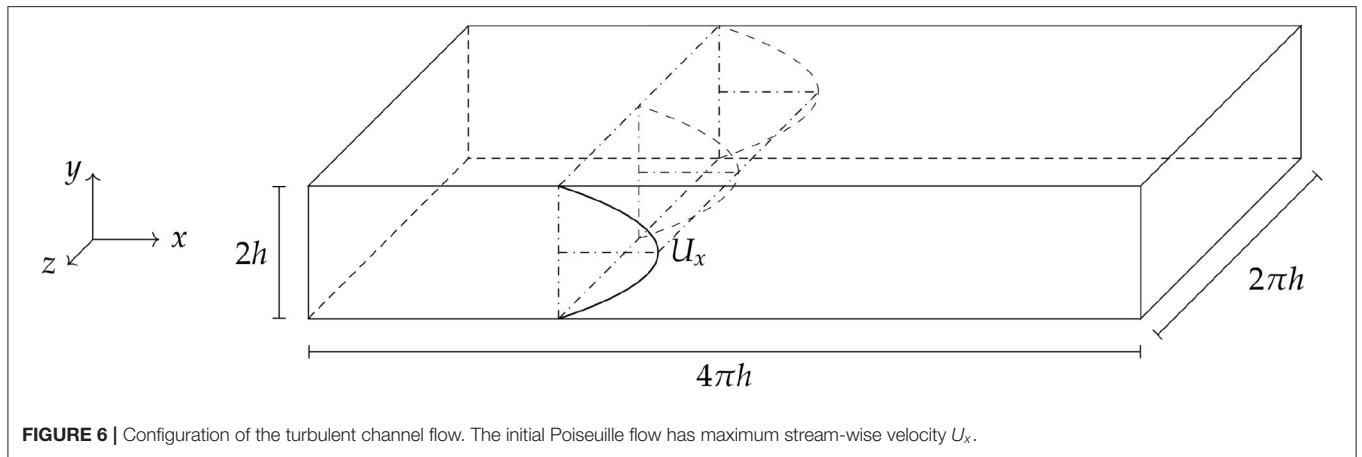


FIGURE 5 | Periodically oscillating flow. Velocity profiles $\langle u_x(\vec{x}, t_{\phi_1}) \rangle_{\phi_1}$ **(A)** and xx -covariance component profiles $\langle u'_x(\vec{x}, t_{\phi_1}) u'_x(\vec{x}, t_{\phi_1}) \rangle_{\phi_1}$ **(B)** from DNS and from DA predictions: along a line parallel to the y -axis crossing the center of the cylinder at $x = 9$ (left); along a line parallel to the y -axis crossing the vortex street at $x = 15$ (right). **(C)** Comparison between the flow field of the ground-truth DNS and of the DA predictions: $\langle u_x(\vec{x}, t_{\phi_1}) \rangle_{\phi_1}$.



windows. The results are in good agreement with the ground truth, even though they are slightly underestimated. The value of the *TKE* computed in the DA predictions are also good for all configurations. The values of *TKE* stabilize around $0.55 \text{ [m}^2/\text{s}^2]$

for all the DA predictions and the more visible differences with respect to the ground truth are in the amplitude of fluctuations. The best agreement is obtained in DA 4. In contrast, the other data windows lead to a loss of accuracy with respect to the ground

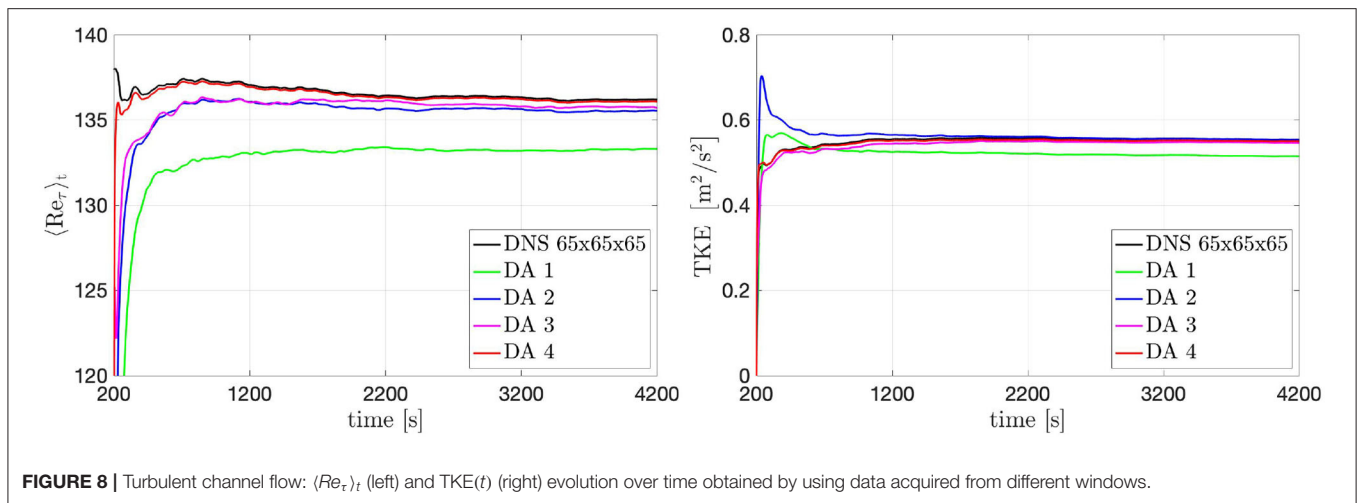


FIGURE 8 | Turbulent channel flow: $\langle Re_\tau \rangle_t$ (left) and TKE(t) (right) evolution over time obtained by using data acquired from different windows.

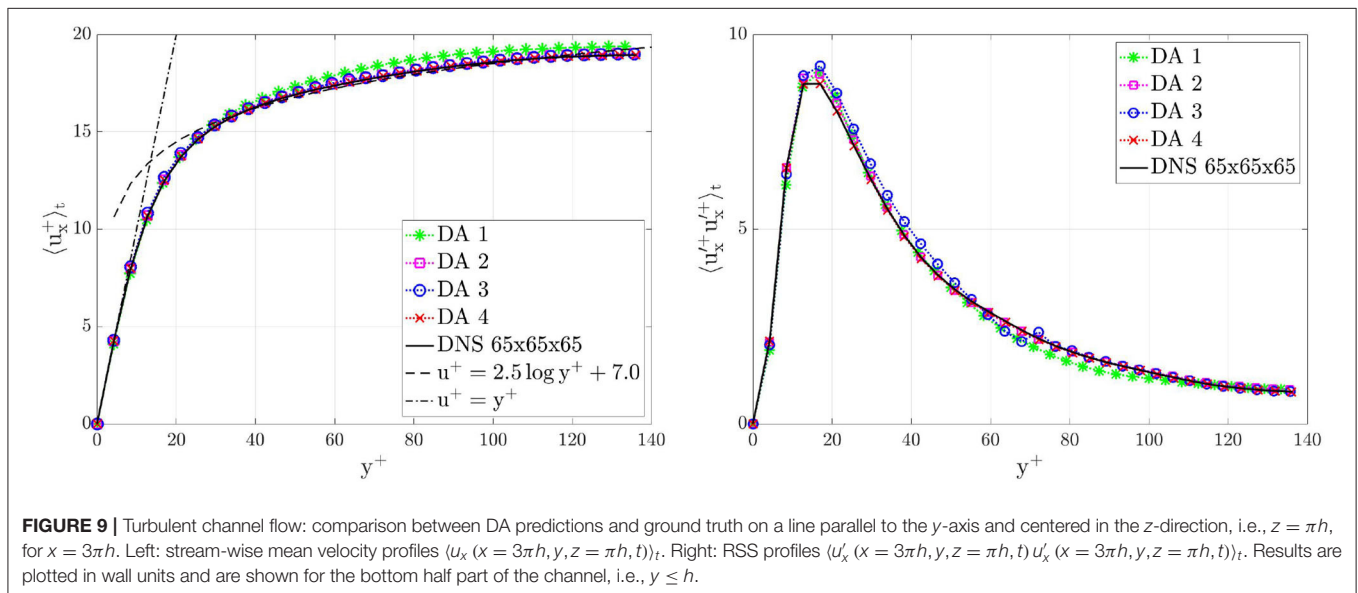


FIGURE 9 | Turbulent channel flow: comparison between DA predictions and ground truth on a line parallel to the y -axis and centered in the z -direction, i.e., $z = \pi h$, for $x = 3\pi h$. Left: stream-wise mean velocity profiles $\langle u_x^+(x = 3\pi h, y, z = \pi h, t) \rangle_t$. Right: RSS profiles $\langle u_x^+(x = 3\pi h, y, z = \pi h, t) u_x^+(x = 3\pi h, y, z = \pi h, t) \rangle_t$. Results are plotted in wall units and are shown for the bottom half part of the channel, i.e., $y \leq h$.

truth. DA 1 returns the least accurate trends, and extension of the window in x - and y -directions, i.e., DA 2 and DA 3, significantly improves accuracy. This means that larger amounts of available data yields DA predictions that tend closer to the ground truth with a shorter numerical transient. For the same reason, DA 4 returns the most accurate predictions.

The time-averaged velocity $\langle u_x^+(\vec{x}, t) \rangle_t$ profiles computed after 4,000 time units are in almost perfect agreement with the ground-truth profile. **Figure 9** shows the comparison between DA prediction and ground truth of the velocity profile $\langle u_x^+(\vec{x}, t) \rangle_t$ and the RSS profile $\langle u_x^+(\vec{x}, t) u_x^+(\vec{x}, t) \rangle_t$ on a line parallel to the y -axis and centered in the z -direction, for $x = 3\pi h$. The DA predictions recover the logarithmic law of the wall $u^+ = A y^+ + B$ (42). The precision slightly decreases for DA 1 where not enough data have been used. Moreover, in the boundary region, the RSS profile $\langle u_x^+(\vec{x}, t) u_x^+(\vec{x}, t) \rangle_t$ present different peaks in the different DA predictions. The more data are used in DA

predictions, the higher the accuracy is. The loss of accuracy for DA 1 is in agreement with the loss of accuracy of the global parameters evolution shown in the **Figure 8**.

3.2.4. Downsampling

In the previous section we evaluated the influence of different window sizes on the accuracy of the results in the DA predictions. Here we want to evaluate the influence of the spatial resolution of the data extracted from the ground-truth solution. To do this, the resolution of the data in window 4 has been down-sampled meaning that the data voxels have a bigger size and the data values are obtained averaging the ground-truth solution over voxels formed by a bigger number of adjacent grid-nodes. The downsampling factor (dw) characterizes the resolution of the data acquisition: for instance, a dw of 2 means that each d_n is obtained averaging the ground-truth solution over voxels

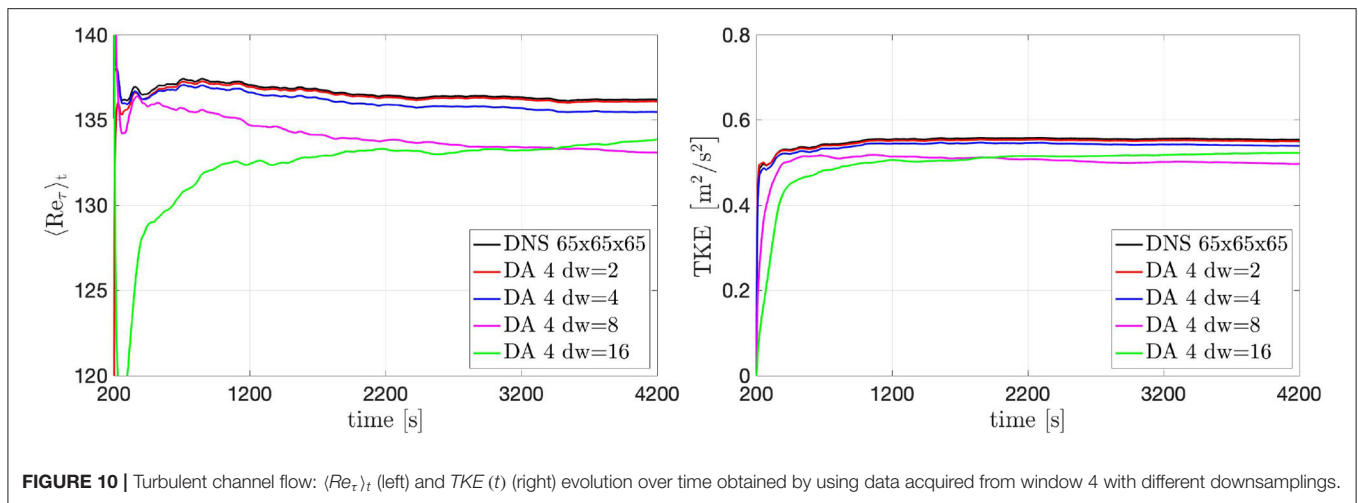


FIGURE 10 | Turbulent channel flow: $\langle Re_{\tau} \rangle_t$ (left) and $TKE(t)$ (right) evolution over time obtained by using data acquired from window 4 with different downsamplings.

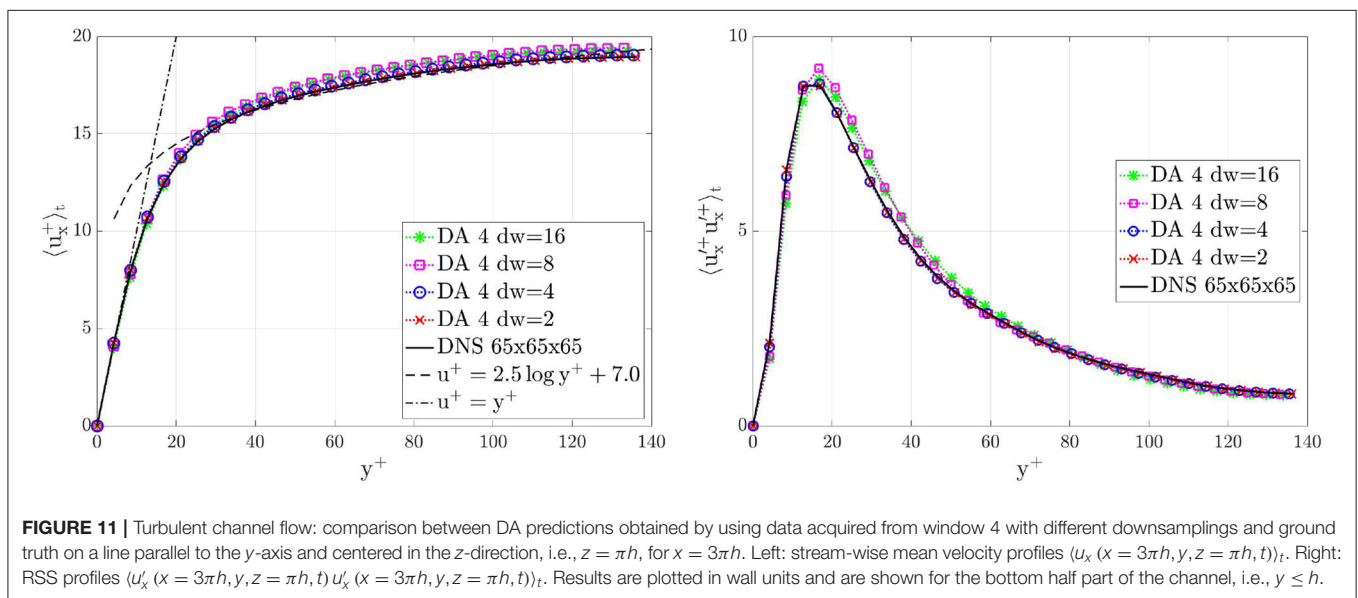


FIGURE 11 | Turbulent channel flow: comparison between DA predictions obtained by using data acquired from window 4 with different downsamplings and ground truth on a line parallel to the y -axis and centered in the z -direction, i.e., $z = \pi h$, for $x = 3\pi h$. Left: stream-wise mean velocity profiles $\langle u_x(x = 3\pi h, y, z = \pi h, t) \rangle_t$. Right: RSS profiles $\langle u_x^+(x = 3\pi h, y, z = \pi h, t) u_x^+(x = 3\pi h, y, z = \pi h, t) \rangle_t$. Results are plotted in wall units and are shown for the bottom half part of the channel, i.e., $y \leq h$.

formed by $(1 + dw) \times (1 + dw) \times (1 + dw)$ adjacent grid-nodes; these voxel-averaged values are then further time-averaged in order to obtain time-and-space averaged data. Here, we analyze the impact of the dw on the results starting from the effect on the $\langle Re_{\tau} \rangle_t$ and $TKE(t)$ evolution, see **Figure 10**. Essentially, a bigger dw (lower data resolution) reduce the accuracy of the solution in the DA predictions: this loose of accuracy is due to the. The results obtained with $dw = 2$ and $dw = 4$ are in a very good agreement with the ground-truth results for both $\langle Re_{\tau} \rangle_t$ and $TKE(t)$.

Figure 11 shows the comparison between DA prediction and ground truth of the velocity profile $\langle u_x^+(\vec{x}, t) \rangle_t$ profile and the RSS profile $\langle u_x^+(\vec{x}, t) u_x^+(\vec{x}, t) \rangle_t$ on a line parallel to the y -axis and centered in the z -direction, computed after 4,000 time units. Downsampling the data increases the magnitude of the covariance \mathbf{r}_n such that the filter will trust in them less than in the forecast solution given by the numerical solver.

The recovery of the log law and the RSS peak predictions slightly deteriorate for higher downsamplings. Even though the unperturbed initial conditions of the flow would not have shown a transition to turbulence without the filter forcing, the boundary conditions are able to maintain the turbulence. After the numerical transient, the loose of accuracy due to the down-sampled data is compensated by the numerical solver. This is an important result because it means that the filter is able to give the correct importance to the two different sources of information.

3.3. Pulsatile Turbulent Flow Downstream a Bioprosthetic Transcatheter Aortic Valve

Here, the configuration setup for a pulsatile turbulent flow past a self-expandable Transcatheter Aortic Valve (TAV) with a heart-rate equal to 70 [beat/min] is investigated in the aortic region

TABLE 1 | Pulsatile turbulent flow downstream a bioprosthetic heart valve: division of the pulse period T in $N_{\phi}^{exp} = 22$ intervals and $N_{\phi}^{num} = 66$ intervals.

ϕ_i^{exp}	1	2	3	4	5	6	7	8	9	10	11	
$t_{\phi_i}^{exp}$ [s]	0.05	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	
ϕ_i^{exp}	12	13	14	15	16	17	18	19	20	21	22	
$t_{\phi_i}^{exp}$ [s]	0.16	0.17	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.60	0.70	
ϕ_i^{exp}	1			2	21			22		
ϕ_i^{num}	1	2	3	4	61	62	63	64	65	66
$t_{\phi_i}^{num}$ [s]	0.05	0.0533	0.0566	0.06	0.60	0.633	0.666	0.70	0.769	0.838

Numerical time intervals have been obtained by refining by a factor of 3 the temporal resolution of each time interval of the experimental data.

downstream the valve. The positive y -direction define the stream-wise direction, while the x - and z -direction are the span-wise ones. The Reynolds number is set to 4,900, based on the nominal diameter $d = 23$ [mm] of the TAV, on the peak (systolic) velocity $U_y = 1$ [m/s] and on the kinematic viscosity ν of the fluid. In this case, no ground truth is available but only experimental data. Moreover, blood flow modeling uncertainty has to be taken into account.

3.3.1. Data Acquisition

The generation of the data $\vec{d}_{v;(s),\phi}$ to be used in the DA predictions has been provided by an experimental setup. The TAV was implanted in a quasi-stiff aortic silicon phantom of the aortic root and integrated in a hydraulic pulse duplicator. Tomographic Particle Image Velocimetry (Tomo-PIV) (43, 44) was used to reconstruct the three-dimensional flow fields in the region of interest. The pulse period $T = 60.0/70$ [s] has been divided in $N_{\phi}^{exp} = 22$ intervals as described in the **Table 1**.

The experimental grid domain has a size of 35 [mm] \times 50 [mm] \times 35 [mm] and consists of $42 \times 59 \times 42$ voxels. Each voxel-and-phase-averaged data has been computed by using 24 repetitions of the heart-pulse, and the covariance of the error of the PIV data results from a phase averaging of instantaneous voxel data fields taken at 24 pulses. Outside the region of interest, the data (velocity and corresponding uncertainty) have been set to zero in order to replicate the presence of the aortic walls. Of course there are additional errors due to the PIV methodology itself which are not considered in the present investigation.

3.3.2. Comparison Between Experiments, Direct Numerical Simulation and Data Assimilation Prediction

Two simulations have been performed for a time equal to $120T$: a DNS where the experimental data are used only to impose inflow and wall boundary conditions; and a DA prediction where the experimental data are used also to filter the forecast flow field by using the DA methodology. The pulse period has been divided in $N_{\phi}^{num} = 66$ intervals obtained by refining by a factor of 3 the temporal resolution of each time interval of the experimental data, as shown in the **Table 1**.

The computational grid used for both DNS and DA prediction has been obtained by refining the experimental grid by a factor of 2 in each direction and then adding an external padding equal to at least 10% to each side of the three-dimensional domain. The final size of the computational grid domain is 44 [mm] \times 60

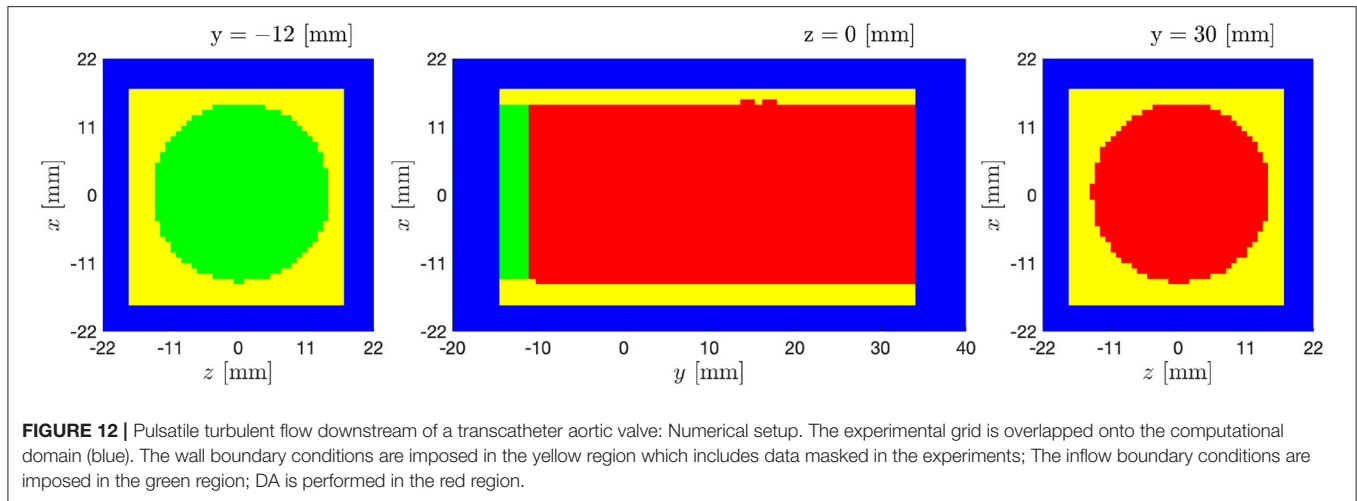
[mm] \times 44 [mm] and consists of $105 \times 145 \times 105$ grid nodes. Mapping the voxel-grid onto the computational grid leads to $g_v = 3 \times 3 \times 3$ grid-nodes lying inside the corresponding voxel. We use periodic boundary conditions for the full computational domain with padding. In both simulations, the presence of the aortic wall and the physical inflow conditions are enforced *via* the fringe forcing method (cf. Section 2.4). The experimental data forced to zero outside the region of interest and the lowest 5 y -plane slices data are used as desired values \vec{U} of the fringe forcing applied in order to impose wall and inflow boundary conditions, respectively. Since the inflow is time dependent, it is not appropriate to use a constant value for λ in the Equation (41). Here, a time-dependent λ has been chosen such that it guarantees the stability condition of the Runge–Kutta time-advancement scheme:

$$\lambda(RK_{substep}, \Delta t) = \frac{1.0}{\alpha(RK_{substep})\Delta t}, \quad (46)$$

where Δt is the numerical time step and α is the first Runge–Kutta coefficient for each Runge–Kutta substep $RK_{substep}$. The inflow boundary conditions for the intermediate numerical time-steps have been imposed by linear interpolation in time of the voxel data. The bulk flow voxel data are used only to perform DA prediction. The initial flow field is set to zero. A sketch of the configuration can be seen in **Figure 12**.

The results obtained by the experiments, DNS and DA prediction have been compared to evaluate the capability of the DA methodology to enhance the spatial and temporal resolution of the observations and to increase the reliability of the characteristic flow patterns captured from the numerical solver. Herein, the systolic phase-averaged velocity $\langle u_y(\vec{x}, t_{\phi}) \rangle_{\phi}$ obtained by the Tomo-PIV experiments, DNS and the DA prediction are shown in **Figures 13A,B** for x -plane and z -plane, respectively.

The DNS and DA prediction present an enhancement in time resolution with respect to the experimental data. The boundary conditions extracted by the experimental data drive the flow in the intermediate iterations between two subsequent applications of the filter. This is an important result because it has been possible to simulate the cardiovascular turbulent flow downstream the heart valve numerically without the need for Fluid-Structure-Interaction methods.



In the DNS, the results show a global overestimation of the velocity field. This is due to boundary condition uncertainty and blood modeling uncertainty. The DA prediction is affected by the same boundary condition and blood modeling uncertainties but the DA prediction strongly reduces the overestimation of the DNS velocity field. It means that the use of the DA methodology has improved the reliability of the velocity field results. The three leaflet jet coming from the bioprosthetic valve seems to be better defined in DA prediction (see, e.g., $t = 0.175T$ slice, in **Figure 13A**). In the lower part the back flow near the left wall, the DA prediction presents a smoother flow field (see **Figure 13B**). This is due to the higher capability of the numerical solver to investigate that region which is, on the contrary, a more difficult challenge for the experimental tool. It is noteworthy that the DA prediction presents a smaller reverse flow region: the filter trusts more in the numerical solver and considers the bigger reverse flow region present in the experiments too uncertain with respect to the forecast solution. Moreover, the DA methodology filters the information coming from those voxel data near the wall that are evidently affected by a tool acquisition error. This error is not visible anymore in DA prediction.

Figure 14A shows the evolution of the flow field for increasingly larger number of pulses performed by the DA algorithm. Velocity profiles from DA prediction, DNS, and experiments are compared in **Figure 14B** along two lines parallel to y -axis. The results show that the flow field prediction has reached a good convergence after 120 pulses.

The finer spatial and time resolution obtained in the DA prediction enhanced the flow field of this specific configuration. This is an important and promising result for future applications of the proposed DA methodology on *in vivo* data from Flow-MRI of patient-specific configurations.

4. DISCUSSION

In the present manuscript, a new DA methodology has been presented. The robust theoretical background of the (S)EnKF

approaches has been applied to pulsatile and turbulent flow configurations. The different ensemble-averaging approaches, that were presented here, show how the study of turbulent flows can be reduced to the definition and the subsequent investigation of the most important statistical properties of the flow, such as mean velocity and the associated fluctuations covariance.

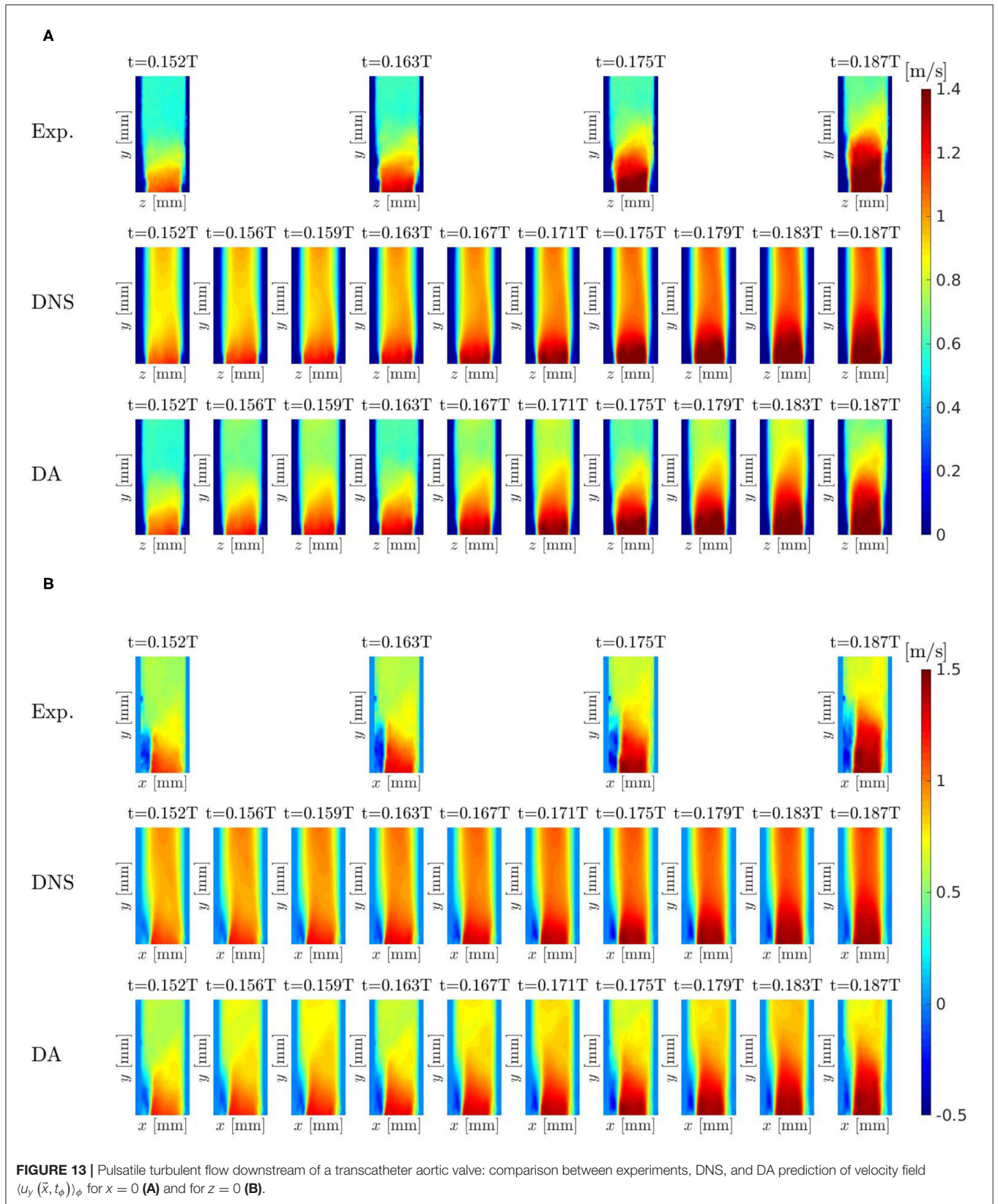
In the context of turbulent flows, the proposed SEnKF-based methodology states as hypothesis that the “true” state-of-system is the mean velocity field. Turbulent fluctuations are deviations from the ensemble-averaged quantities. This leads to the formal equality between turbulent fluctuations and the observation error \tilde{r}_n required in the derivation of a SEnKF.

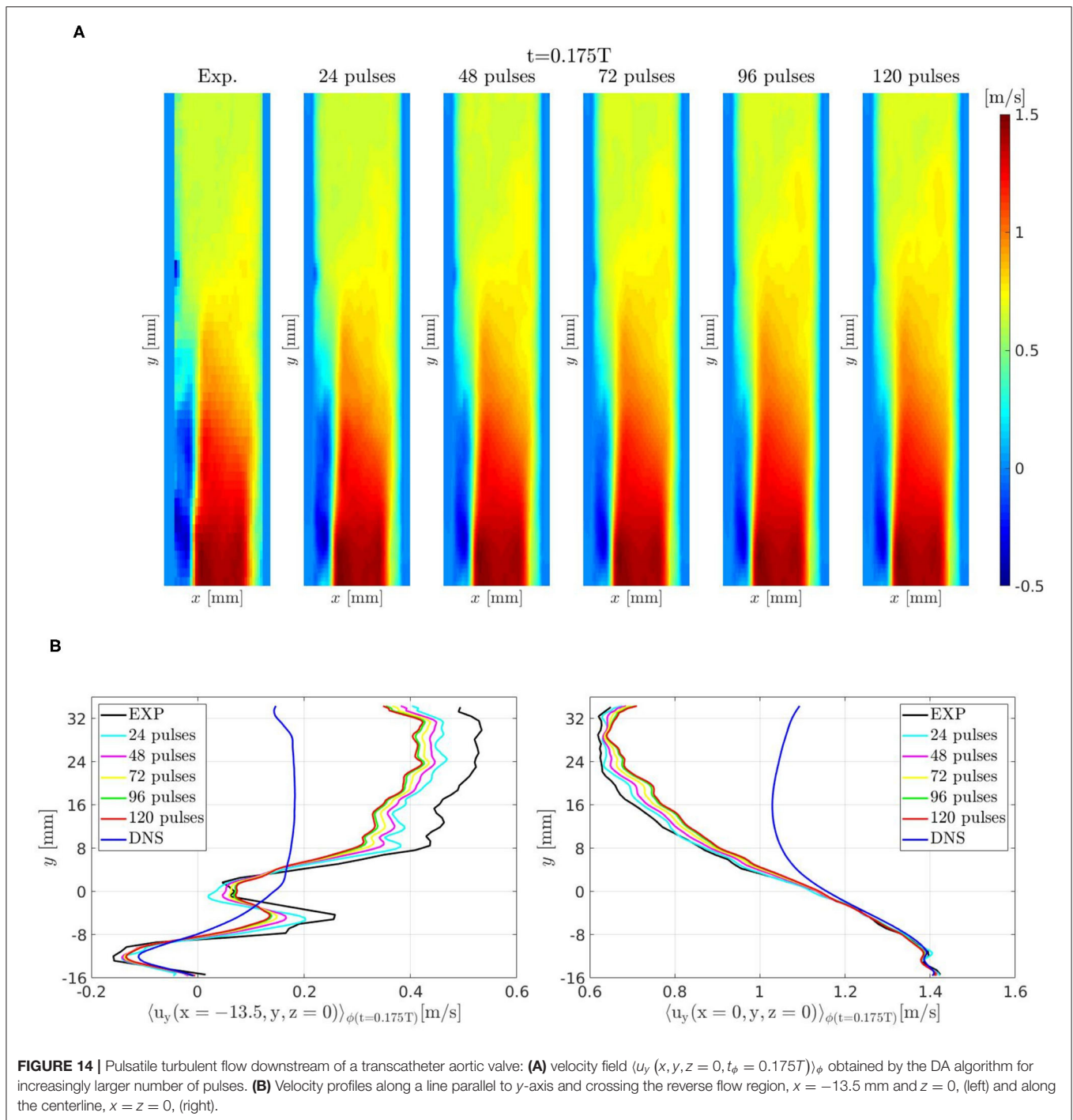
In the derivation of our DA methodology, the definition of the states of the ensemble allows us to treat the forecast error \tilde{q}_n in the proper way: this error has zero mean and the contribution of the covariance of the numerical error $\tilde{q}_{n_{NUM}}$ will be negligible compared with the one of the turbulent error $\tilde{q}_{n_{TUR}}$ and modeling error $\tilde{q}_{n_{MOD}}$. In the first two test cases (sections 3.1, 3.2), the model is well defined and therefore $\tilde{q}_{n_{MOD}}$ is zero. In the last test case (section 3.3), DNS and DA predictions are affected by the same $\tilde{q}_{n_{MOD}}$. Therefore, only the turbulent error is relevant in the present study.

The choice of the ensemble of states-of-system in the case of pulsatile and turbulent flows allows one to filter the forecast solution at each time t_n with the corresponding data flow field; this yields to an additional but constant computational cost for each pulse interval.

From a numerical point of view, filtering the forecast solution with a forcing term only in nonconsecutive time-step iterations could be seen as forcing the flow field with a temporal Dirac δ function. This would lead to numerical oscillations in the intermediate iterations between two filtering iterations and to a local non-smooth advancement in time of the numerical solution.

First, the DA methodology has been validated vs. an unsteady flow past a cylinder confined in a channel using data acquired





from previous numerical simulations. In the DA predictions, the presence of the cylinder was not taken into account by using any boundary conditions but nevertheless the reconstruction of the wake downstream the cylinder has been successfully reconstructed by filtering the forecast solution with the data. The test shows that the DA method, and in particular the proposed choice of the ensemble of states-of-system, copes with unsteady flow field presenting a periodicity in time.

Second, the methodology was extended to turbulent flow applications. The statistically-steady configuration of the wall-bounded turbulent flow shows that the turbulent features can be reconstructed by using the data even though the numerical simulation is initialized with a laminar flow which would not exhibit a transition to turbulence by itself. The effect of the different data configurations on the bulk flow and on the walls has been investigated. Downsampling the resolution of data

acquisition leads to a loss of accuracy in the DA prediction that is limited to 5%.

Finally, the methodology has been applied to pulsatile and turbulent configuration of a flow past a bioprosthetic heart valve. The data of a pulsatile turbulent flow past a TAV, acquired by a Tomo-PIV technique have been used to reconstruct the wake in the aorta in a computational simulation with higher spatial and temporal resolution. The test-case shows how the DA methodology deals with FSI applications even though no structural modeling is required. The results show the capability, the robustness and the accuracy of the proposed methodology to cope with realistic configurations of biomedical applications.

The presented results show that the filter is capable of interpreting the accuracy and reliability of the data with respect to the numerical solution giving locally a different weight to the two different sources of information. A bigger amount of data, although with greater uncertainty, allows to lead the solution of the DA method closer to the true-state of the system. This is useful in heart valve applications where the data acquired by experimental investigations or through 4D Flow MRI techniques suffer from lower accuracy near the walls; these noisy data are filtered and the higher accuracy of the numerical solver will return a better description of the flow field in such regions. On the other hand the available data of the specific-patient configuration will increase the accuracy of the numerical solution in the bulk flow if compared with the solution obtained by using a FSI solver employing a simplified morphology of the aortic valve and aortic root.

A first limitation of the methodology is that we consider the covariance across different voxels to be zero. This is equivalent of assuming that the autocorrelation length is less than the voxel size h . This limitation might be overcome solving with higher computational cost the analysis step by using the global matrices \mathbf{H} , \mathbf{P}^f and \mathbf{R} instead of \mathbf{h} , \mathbf{p}^f and \mathbf{r} . A second limitation of the methodology is represented by the definition of the matrix \mathbf{D} which is required to ensure positiveness of covariance matrix \mathbf{p}^f . From a theoretical point of view \mathbf{D} influence disappears for increasingly large ensembles, but the choice of its value strongly affects the convergence history of the DA prediction. For \mathbf{D} too high, the DA prediction will require a larger number of cycles in order to reach a converged solution for both mean velocity and covariance. Moreover, a large value of \mathbf{D} reduces the physical meaning of the matrix \mathbf{p}^f and this will influence the sensitivity of the filter. On the other hand, for \mathbf{D} too low the forecast solution would be considered closer to the “true” state-of-system by the filter even though it cannot be stated a priori. In conclusion, the value of \mathbf{D} has to be chosen high and run the DA prediction for the required number of cycles until convergence is reached. For this reason, in the DA prediction of the pulsatile turbulent flow downstream a bioprosthetic heart valve, 120 pulses (5 times the number of repetitions available by the experiments) have been enough to enhance the description of the mean flow fields with respect to the observations, but have not produced a significant improvement of the RSS fields.

The investigation of the RSS fields requires to collect a larger ensemble for both matching the classical requirements on the ensemble size of turbulent statistics and fully overcoming the influence of \mathbf{D} . The sensitivity of the proposed methodology to the matrix \mathbf{D} and the size of ensemble needed for convergence in these cardiovascular applications in the ascending aorta will be further investigated in order to estimate the computational cost required to enhance RSS fields. A further limitation is that synthetic experimental data have been created without adding any Gaussian error after spatially averaging inside the voxel, even though this has been previously done in literature. This choice has been made here, because we want to focus our investigations on the effect of turbulent fluctuations (considered as measurement noise). Additional noise may pose additional problems to the described method, in particular, because the real noise in MRI data has a non-Gaussian character. In practice, 4D flow MRI quantification would require more complex methods in order to quantify the effects of these hardware source of errors. Specific methods have been developed for creating synthetic 4D flow MRI data from raw Phase Contrast MRI data to better assess turbulent features, e.g., turbulence intensity and TKE (45–50), and for completely excluding the effects of the hardware source errors by generating synthetic MRI data fields (51).

In conclusion, the results show that the method is promising for future use with *in vivo* data from 4D Flow MRI.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

DD and DO wrote the manuscript and reviewed the manuscript and designed the data assimilation method. DD implemented the data assimilation method in the numerical solver and performed the simulations. DO supervised the scientific work. Both authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the Platform for Advanced Scientific Computing (PASC, <http://www.pasc-ch.org>) under the HPC-PREDICT project (<https://www.pasc-ch.org/projects/2017-2020/hpc-predict>).

ACKNOWLEDGMENTS

The authors acknowledge Valérie Kulka and Derick Sivakumaran who contributed in developing post-processing tools; David Hasler and Leonardo Pietrasanta who provided the Tomo-PIV experimental data. The authors also acknowledge the support of the Swiss National Supercomputing Center (CSCS, <http://www.cscs.ch>).

REFERENCES

- Binter CE, Gotschy A, Sündermann SH, Frank M, Tanner FC, Lüscher TF, et al. Turbulent kinetic energy assessed by multipoint 4-dimensional flow magnetic resonance imaging provides additional information relative to echocardiography for the determination of aortic stenosis severity. *Circ Cardiovasc Imaging*. (2017) 10:e005486. doi: 10.1161/CIRCIMAGING.116.005486
- Kalman RE. A new approach to linear filtering and prediction problems. *J Basic Eng*. (1960) 82:35–45. doi: 10.1115/1.3662552
- Sorenson HW. *Kalman Filtering: Theory and Application*. IEEE Press (1985).
- Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res*. (1994) 99:10143–62. doi: 10.1029/94JC00572
- Burgers G, Jan van Leeuwen P, Evensen G. Analysis scheme in the ensemble Kalman filter. *Mon Weather Rev*. (1998) 126:1719–24. doi: 10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2
- Houtekamer PL, Mitchell HL. Data assimilation using an ensemble Kalman filter technique. *Mon Weather Rev*. (1998) 126:796–811. doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2
- Höpfner J, Chevalier M, Bewley TR, Henningson DS. State estimation in wall-bounded flow systems. Part 1. Perturbed laminar flows. *J Fluid Mech*. (2005) 534:263–94. doi: 10.1017/S0022112005004210
- Chevalier M, Höpfner J, Bewley TR, Henningson DS. State estimation in wall-bounded flow systems. Part 2. Turbulent flows. *J Fluid Mech*. (2006) 552:167–87. doi: 10.1017/S0022112005008578
- Colburn C, Cessna J, Bewley T. State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter. *J Fluid Mech*. (2011) 682:289–303. doi: 10.1017/jfm.2011.222
- Gu Y, Oliver DS. An iterative ensemble Kalman filter for multiphase fluid flow data assimilation. *SPE J*. (2007) 12:438–46. doi: 10.2118/108438-PA
- Harlim J, Majda AJ. Filtering turbulent sparsely observed geophysical flows. *Mon Weather Rev*. (2010) 138:1050–83. doi: 10.1175/2009MWR3113.1
- Suzuki T. Reduced-order Kalman-filtered hybrid simulation combining particle tracking velocimetry and direct numerical simulation. *J Fluid Mech*. (2012) 709:249–88. doi: 10.1017/jfm.2012.334
- Meldi M, Poux A. A reduced order model based on Kalman filtering for sequential data assimilation of turbulent flows. *J Comput Phys*. (2017) 347:207–34. doi: 10.1016/j.jcp.2017.06.042
- Foures DP, Dovetta N, Sipp D, Schmid PJ. A data-assimilation method for Reynolds-averaged Navier-Stokes-driven mean flow reconstruction. *J Fluid Mech*. (2014) 759:404–31. doi: 10.1017/jfm.2014.566
- Symon S, Dovetta N, McKeon BJ, Sipp D, Schmid PJ. Data assimilation of mean velocity from 2D PIV measurements of flow over an idealized airfoil. *Exp Fluids*. (2017) 58:61. doi: 10.1007/s00348-017-2336-8
- Mons V, Chassaing JC, Gomez T, Sagaut P. Reconstruction of unsteady viscous flows using data assimilation schemes. *J Comput Phys*. (2016) 316:255–80. doi: 10.1016/j.jcp.2016.04.022
- Gaidzik F, Stucht D, Roloff C, Speck O, Thévenin D, Janiga G. Transient flow prediction in an idealized aneurysm geometry using data assimilation. *Comput Biol Med*. (2019) 115:103507. doi: 10.1016/j.compbiomed.2019.103507
- Canuto D, Pantoja JL, Han J, Dutson EP, Eldredge JD. An ensemble Kalman filter approach to parameter estimation for patient-specific cardiovascular flow modeling. *Theor Comput Fluid Dyn*. (2020) 34:521–44. doi: 10.1007/s00162-020-00530-2
- DeVault K, Gremaud PA, Novak V, Olufsen MS, Vernieres G, Zhao P. Blood flow in the circle of willis: modeling and calibration. *Multiscale Model Simulat*. (2008) 7:888–909. doi: 10.1137/07070231X
- Arnold A, Battista C, Bia D, German YZ, Armentano RL, Tran H, et al. Uncertainty quantification in a patient-specific one-dimensional arterial network model: EnKF-based inflow estimator. *J Verificat Validat Uncert Quant*. (2017) 2:011002. doi: 10.1115/1.4035918
- Habibi M, D'Souza RM, Dawson ST, Arzani A. Integrating multi-fidelity blood flow data with reduced-order data assimilation. *Comput Biol Med*. (2021). 135:104566. doi: 10.1016/j.compbiomed.2021.104566
- Arzani A, Dawson ST. Data-driven cardiovascular flow modelling: examples and opportunities. *J R Soc Interface*. (2021) 18:20200802. doi: 10.1098/rsif.2020.0802
- Rispoli VC, Nielsen JF, Nayak KS, Carvalho JL. Computational fluid dynamics simulations of blood flow regularized by 3D phase contrast MRI. *Biomed Eng Online*. (2015) 14:1–23. doi: 10.1186/s12938-015-0104-7
- Koltukluoglu TS, Blanco PJ. Boundary control in computational haemodynamics. *J Fluid Mech*. (2018) 847:329–64. doi: 10.1017/jfm.2018.329
- Fathi MF, Bakhshinejad A, Baghaie A, Saloner D, Sacho RH, Rayz VL, et al. Denoising and spatial resolution enhancement of 4D flow MRI using proper orthogonal decomposition and lasso regularization. *Comput Med Imaging Graphics*. (2018) 70:165–72. doi: 10.1016/j.compmedimag.2018.07.003
- Annio G, Torii R, Ariff B, O'Regan DP, Muthurangu V, Ducci A, et al. Enhancing magnetic resonance imaging with computational fluid dynamics. *J Eng Sci Med Diagn Therapy*. (2019) 2:041010. doi: 10.1115/1.4045493
- Ferdian E, Suinesiaputra A, Dubowitz DJ, Zhao D, Wang A, Cowan B, et al. 4DFlowNet: super-resolution 4D flow MRI using deep learning and computational fluid dynamics. *Front Phys*. (2020). 8:138. doi: 10.3389/fphys.2020.00138
- Lewis JM, Derber JC. The use of adjoint equations to solve a variational adjustment problem with advective constraints. *Tellus A*. (1985) 37:309–322. doi: 10.3402/tellusa.v37i4.11675
- Le Dimet FX, Talagrand O. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*. (1986) 38:97–110. doi: 10.3402/tellusa.v38i2.11706
- Evensen G, Van Leeuwen PJ. An ensemble Kalman smoother for nonlinear dynamics. *Mon Weather Rev*. (2000) 128:1852–67. doi: 10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2
- Henniger R, Obrist D, Kleiser L. High-order accurate solution of the incompressible Navier-Stokes equations on massively parallel computers. *J Comput Phys*. (2010) 229:3543–72. doi: 10.1016/j.jcp.2010.01.015
- Henniger R, Kleiser L, Meiburg E. Direct numerical simulations of particle transport in a model estuary. *J Turbulence*. (2010) 11:N39. doi: 10.1080/14685248.2010.516257
- John MO, Obrist D, Kleiser L. Stabilisation of subcritical bypass transition in the leading-edge boundary layer by suction. *J Turbulence*. (2014) 15:795–805. doi: 10.1080/14685248.2014.933226
- Burns P, Meiburg E. Sediment-laden fresh water above salt water: nonlinear simulations. *J Fluid Mech*. (2015) 762:156. doi: 10.1017/jfm.2014.645
- Nestola MGC, Becsek B, Zolfaghari H, Zulian P, De Marinis D, Krause R, et al. An immersed boundary method for fluid-structure interaction based on variational transfer. *J Comput Phys*. (2019) 398:108884. doi: 10.1016/j.jcp.2019.108884
- Becsek B, Pietrasanta L, Obrist D. Turbulent systolic flow downstream of a bioprosthetic aortic valve: velocity spectra, wall shear stresses, and turbulent dissipation rates. *Front Physiol*. (2020) 11:577188. doi: 10.3389/fphys.2020.577188
- Nordström J, Nordin N, Henningson D. The fringe region technique and the fourier method used in the direct numerical simulation of spatially evolving viscous flows. *SIAM J Sci Comput*. (1999) 20:1365–93. doi: 10.1137/S1064827596310251
- Kim J, Moin W, Moser R. Turbulence statistics in fully developed channel flow at low Reynolds number. *J Fluid Mech*. (1987) 177:133–66. doi: 10.1017/S0022112087000892
- Pedley TJ. *The Fluid Mechanics of Large Blood Vessels*. Cambridge: Cambridge Monographs on Mechanics; Cambridge University Press (1980).
- Schlatter P, Stolz S, Kleiser L. Evaluation of high-pass filtered eddy-viscosity models for large-eddy simulation of turbulent flows. *J Turbulence*. (2005) 6:N5. doi: 10.1080/14685240500103127
- Schmid P, Henningson DS. *Stability and Transition in Shear Flows*. New York, NY: Springer (2000).
- Pope SB. *Turbulent Flows*. Cambridge: Cambridge Univ. Press (2011).
- Hasler D, Landolt A, Obrist D. Tomographic PIV behind a prosthetic heart valve. *Exp Fluids*. (2016) 57:80. doi: 10.1007/s00348-016-2158-0
- Hasler D, Obrist D. Three-dimensional flow structures past a bio-prosthetic valve in an *in-vitro* model of the aortic root. *PLoS ONE*. (2018) 13:e0194384. doi: 10.1371/journal.pone.0194384

45. Ha H, Lantz J, Haraldsson H, Casas B, Ziegler M, Karlsson M, et al. Assessment of turbulent viscous stress using ICOSA 4D Flow MRI for prediction of hemodynamic blood damage. *Sci Rep.* (2016) 6:1–14. doi: 10.1038/srep39773
46. Ziegler M, Lantz J, Ebberts T, Dyverfeldt P. Assessment of turbulent flow effects on the vessel wall using four-dimensional flow MRI. *Magn Reson Med.* (2017) 77:2310–19. doi: 10.1002/mrm.26308
47. Ha H, Kvitting JP, Dyverfeldt P, Ebberts T. Validation of pressure drop assessment using 4D flow MRI-based turbulence production in various shapes of aortic stenoses. *Magn Reson Med.* (2019) 81:893–906. doi: 10.1002/mrm.27437
48. Ha H, Kvitting JPE, Dyverfeldt P, Ebberts T. 4D Flow MRI quantification of blood flow patterns, turbulence and pressure drop in normal and stenotic prosthetic heart valves. *Magn Reson Imag.* (2019). 55:118–27. doi: 10.1016/j.mri.2018.09.024
49. Walheim J, Dillinger H, Gotschy A, Kozerke S. 5D flow tensor MRI to efficiently map Reynolds stresses of aortic blood flow *in-vivo*. *Sci Rep.* (2019). 9:1–12. doi: 10.1038/s41598-019-55353-x
50. Walheim J, Dillinger H, Kozerke S. Multipoint 5D flow cardiovascular magnetic resonance-accelerated cardiac-and respiratory-motion resolved mapping of mean and turbulent velocities. *J Cardiovasc Magn Reson.* (2019) 21:1–13. doi: 10.1186/s12968-019-0549-0
51. Puiseux T, Sewonu A, Moreno R, Mendez S, Nicoud F. Numerical simulation of time-resolved 3D phase-contrast magnetic resonance imaging. *PLoS ONE.* (2021) 16:e0248816. doi: 10.1371/journal.pone.0248816

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 De Marinis and Obrist. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.