12th International Symposium on Turbulence and Shear Flow Phenomena (TSFP12)
Osaka, Japan (Online), July 19-22, 2022

COMPLETE FLOW CHARACTERIZATION FROM SNAPSHOT PIV, FAST PROBES AND PHYSICS-INFORMED NEURAL NETWORKS

Álvaro Moreno Soto, Alejandro Güemes and Stefano Discetti
Aerospace Engineering Research Group
Universidad Carlos III de Madrid
Avda. de la Universidad 30, Leganés 28911, Spain
amsoto@ing.uc3m.es

ABSTRACT

The use of physics-informed neural networks (PINNs), which incorporate governing laws to constrain the training of machine-learning algorithms, has opened the door to novel applications of artificial intelligence (AI) to model and regularize experimental data. PINNs have been recently shown to improve the accuracy of time-resolved measurements, but its capabilities are significantly reduced when that resolution is not at hand. In this work, we exploit PINNs to enhance velocity measurements from non-time-resolved field measurements, as those available for example in snapshot Particle Image Velocimetry (PIV). We propose using PINNs as a regularizer of time-resolved estimated fields from simultaneous measurements with fast pointwise probes and non-time-resolved PIV. We use a multilayer perceptron (MLP) architecture to establish a correspondence between probe data and the temporal coefficients of the Proper Orthogonal Decomposition of the velocity profiles. The estimated fields are then fed to the PINNs to enhance the data and increase precision, additionally extracting derived quantities not available in the raw data, such as the pressure distribution.

INTRODUCTION

Machine-learning (ML) algorithms are recently being embedded in an ever-growing amount of applications due to both the increase in the computational capacity as well as the speed at which results can be obtained. Among others applications, ML has demonstrated being a powerful ally to recover, correct, regularize or enhance (among other capabilities) experimental data that otherwise would not be accessible due to setup or hardware limitations. Physics-informed neural networks (PINNs) (Raissi *et al.* (2019)) are an example of function approximator based on a neural network architecture trained by directly enforcing physical constraints. In their original implementation, PINNs were based on fully-connected layers which incorporate governing laws as loss function and therefore are able to impose physics constraints while reconstructing field variables from incomplete or limited measurements or data (Mao *et al.* (2020); Cai *et al.* (2021); Yang *et al.* (2021)).

In the field of fluid mechanics, PINNs have been proven to provide promising results regarding fluid flow estimation (Raissi *et al.* (2020)) using scalar field data. For this case in particular, the normalized field variables, namely the components of the velocity vector in 3D (u,v,w) and the pressure p, follow the well-known Navier-Stokes (NS) equations, which can be written in a simplified manner as follows:

$$\begin{cases} \frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \\ \frac{\partial u}{\partial t}+\left(u\frac{\partial u}{\partial x}+v\frac{\partial u}{\partial y}+w\frac{\partial u}{\partial z}\right)=-\frac{\partial p}{\partial x}+\frac{1}{Re}\left(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}\right) \\ \frac{\partial v}{\partial t}+\left(u\frac{\partial v}{\partial x}+v\frac{\partial v}{\partial y}+w\frac{\partial v}{\partial z}\right)=-\frac{\partial p}{\partial y}+\frac{1}{Re}\left(\frac{\partial^2 v}{\partial x^2}+\frac{\partial^2 v}{\partial y^2}+\frac{\partial^2 v}{\partial z^2}\right) \\ \frac{\partial w}{\partial t}+\left(u\frac{\partial w}{\partial x}+v\frac{\partial w}{\partial y}+w\frac{\partial w}{\partial z}\right)=-\frac{\partial p}{\partial z}+\frac{1}{Re}\left(\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}+\frac{\partial^2 w}{\partial z^2}\right), \end{cases} \tag{1}$$

where t,x,y,z refer to time and Cartesian coordinates, and $Re=UL/\nu$ is Reynolds number based on a characteristic velocity U, length L and kinematic viscosity ν. This set of equations is enforced as a constraint to reconstruct and regularize fluid velocity fields.

PINNs have been recently used to obtain dense velocity fields from Lagrangian Particle Tracking measurements (Han *et al.* (2021); Wang *et al.* (2022)). The implementation of PINNs for data assimilation using sparse particle tracks requires time-resolved data to accurately compute the time derivatives in equation 1. This is often difficult to achieve due to hardware limitations, and when feasible it requires in most cases expensive equipment. This limits the use of PINNs for non-time-resolved Particle Image Velocimetry (PIV), often referred as 'snapshot PIV'.

To overcome the lack of time resolution, several studies have aimed at predicting flow fields via multi-time delay linear estimators (and more recently non-linear mapping based on neural networks) using pointwise probes, for which time resolution is more accessible, in combination with snapshot PIV. While this approach comes at the expense of increased complexity of the setup, it has demonstrated being an interesting solution to obtain time-resolved flow field measurements. Estimators based on Linear Stochastic Estimation (see for instance Tinney *et al.* (2008) for jet flows), Extended Proper Orthogonal Decomposition (Borée, 2003) and more recently recurrent neural networks (Deng *et al.*, 2019; Jin *et al.*, 2020) have been proven effective to obtain good estimations of the most energetic flow features. On the downside, flow field features exhibiting low level of correlation with the probe data were reconstructed with higher error. Furthermore, in most cases data were mapped onto a lower rank space to improve the identification of correlated features, thus inevitably introducing truncation error into the reconstructed fields.

In this work, we present a novel architecture including PINNs as a regularizer for time-resolved flow fields estimated from fast pointwise probes. This approach features several benefits. First, we enable the use of PINNs for non-time-resolved data by introducing time-resolution through the combination of PIV with probes. Second, PINNs are expected to improve the accuracy of the reconstructed fields by imposing

physical constraints. Third, with adequate boundary conditions PINNs can augment the data by providing access to additional quantities, such as the pressure field.

In the Methodology section, the flow chart of the algorithm is detailed. In the Validation and Results section, the test case is described and the main results are outlined, respectively. Finally, the conclusions are drawn.

METHODOLOGY

The proposed algorithm is sketched in figure 1. We assume snapshot PIV data are available on a regular grid. In the sketch the input data are shown for simplicity in a planar domain, although a fully rigorous implementation would require 3D velocity measurements.

Firstly, a Proper Orthogonal Decomposition (POD) of the velocity fields is carried out to identify a low-rank space to approximate the velocity data. This operation is carried out through the snapshot POD implementation (Sirovich (1987)), corresponding to the economy-size Singular Value Decomposition:

$$U = \Psi \Sigma \Phi^T \tag{2}$$

In the above equation, Ψ and Φ are square unitary matrices containing respectively the temporal and spatial modes, and Σ is a diagonal matrix whose elements are the singular values.

To obtain time-resolved measurements, we need to collect data with time resolution from point probes. Practical limitations of setup complexity limit the maximum number of probes (see for instance Discetti *et al.* (2019), using 5 probes for time-resolved flow estimation in a high-Reynolds-number pipe flow). In order to increase the available information to reconstruct each snapshot, in convection-dominated flow the temporal information can be directly embedded to generate virtual probes. This multi-time delay approach has been used by several authors (Tinney *et al.*, 2008; Hosseini *et al.*, 2015; Discetti *et al.*, 2018) and is adopted also here to enhance the probe information to reconstruct fields.

The probe data are then fed as input to a multilayer perceptron (MLP), with the output being the time coefficients of the most energetic temporal POD modes of the velocity fields. Once the model is trained on a set of probe and PIV data captured simultaneously, it can encode the temporal modes of the velocity field at any other time given the information provided by the probes. Consequently, we are able to reconstruct an approximate time-resolved $\tilde{\Psi}_{TR}$ via the input of other subsequent velocity data points at any given time, recovering from non-time-resolved PIV information a time-resolved approximation of the temporal modes of the target velocity field. The reason for using POD as a field encoder resides in its simplicity of embedding high-dimensional information (3D PIV data can easily provide 100k vectors per snapshot) in a lower rank space, which simplifies the learning process and eases the convergence. It must be remarked that the naive approach used here to embed the time information can be overcome by using, for instance, long short-term memory neural networks (Hochreiter & Schmidhuber (1997); Kalchbrenner *et al.* (2015); Greff *et al.* (2017)). Here, the use of MLP is motivated by its simplicity of implementation and its integration with the PINNs architecture, although in future works recurrent architectures may be considered.

Once this approximate time-resolved matrix $\tilde{\Psi}_{TR}$ is obtained, the full velocity field may be recovered on the same grid as the original PIV via $\tilde{\mathscr{U}}_{TR} = \tilde{\Psi}_{TR}\Sigma_{PIV}\Phi^T_{PIV}$. The approximated time-resolved velocity field $\tilde{\mathscr{U}}_{TR}$ is now resolved both in space and time, but lacks precision due to accumulation of errors from several sources, including above all POD truncation and inaccuracies in the temporal-mode estimation with the MLP. To compensate for those errors and regularize the estimated fields, PINNs are used to improve the accuracy of the fluid domain with the regularization being based on the physical constraint of equation 1. Including boundary conditions, PINNs can also augment the data disclosing additional quantities, such as the pressure field, which was not available at first.

The training of the MLP and of the PINN is carried out separately. For the MLP, a simple cost function based on the minimization of the error in the reconstruction of the temporal modes is implemented. Regarding the PINN, the loss function $\mathscr{L}$ contains three main contributions:

- the residual of the Navier-Stokes equations, $\mathscr{L}_{NS}$, computed on a target refined grid to reduce truncation errors,
- the error with respect to the reference PIV velocity data on the PIV grid, $\mathscr{L}_{PIV}$,
- and, if available, a boundary condition on pressure at any given point or set of points, $\mathscr{L}_p$. This boundary condition could arise from direct measurement of the pressure in points of the domain (not necessarily the boundaries) and / or from reasonable assumptions on the boundaries of the domain.

The residual of the NS equations is computed by the mean squared error (mse) of each equation independently (continuity and two momentum equations for a 2D configuration such as the case presented in this article), i.e. $\mathscr{L}_{NS} = \text{mse}(0,e_1) + \text{mse}(0,e_2) + \text{mse}(0,e_3)$, where 0 would be the reference residual value if NS were to be perfectly obeyed and e_1, e_2 and e_3 the residuals of NS from the predicted model, correspondingly. Similarly, $\mathscr{L}_{PIV}$ and $\mathscr{L}_p$ are also calculated via mse between the predicted quantities by the model and the reference PIV and pressure values, respectively. It must be remarked that the PIV data might be also affected by uncertainty, i.e. in principle, even in case of perfect assimilation and exact estimation of the velocity field, the loss function would not be zero.

In previous studies the loss function is computed as the sum of the three losses. In this work, we decided to opt for an adaptative-weighted error function, i.e. the weight of each error contribution is tuned during the training process. More specifically, the error with the higher value is adapted with the higher weight following a direct proportionality rule, i.e. $\mathscr{L} = w_{NS}\mathscr{L}_{NS} + w_{PIV}\mathscr{L}_{PIV} + w_p\mathscr{L}_p$, where $w_\bullet = \mathscr{L}_\bullet/(\mathscr{L}_{NS} + \mathscr{L}_{PIV} + \mathscr{L}_p)$. Therefore, the total loss may be written in a simplified manner as $\mathscr{L} = (\mathscr{L}^2_{NS} + \mathscr{L}^2_{PIV} + \mathscr{L}^2_p)/(\mathscr{L}_{NS} + \mathscr{L}_{PIV} + \mathscr{L}_p)$. This approach is justified by the need to compensate between two main counteractive effects during training: on the first hand, enforcing the compliance of NS tends to homogenize the system, whereas the input of reference values imposes that the solution is in accordance with experimentally-accessible data (which, as discussed above, are not necessarily accurate). In the authors' experience, in the first part of the training process the effort is more dedicated to match the estimation with the experimental data (i.e. minimizing $\mathscr{L}_{PIV}$). As $\mathscr{L}_{PIV}$ is reduced, $\mathscr{L}_{NS}$ gains progressively relative importance and the physical constraint regularizes the data.

VALIDATION DATASET AND SETTINGS

To validate our proposed architecture, we make use of synthetic data calculated via Direct Numerical Simulation (DNS) of a fluid flow surrounding a pinball configuration (Deng *et al.*, 2020), i.e. three cylinders of diameter D located at the vertexes of an equilateral triangle with lateral distance $l = 3/2D$. The simulation consists of 20,000 snapshots (time

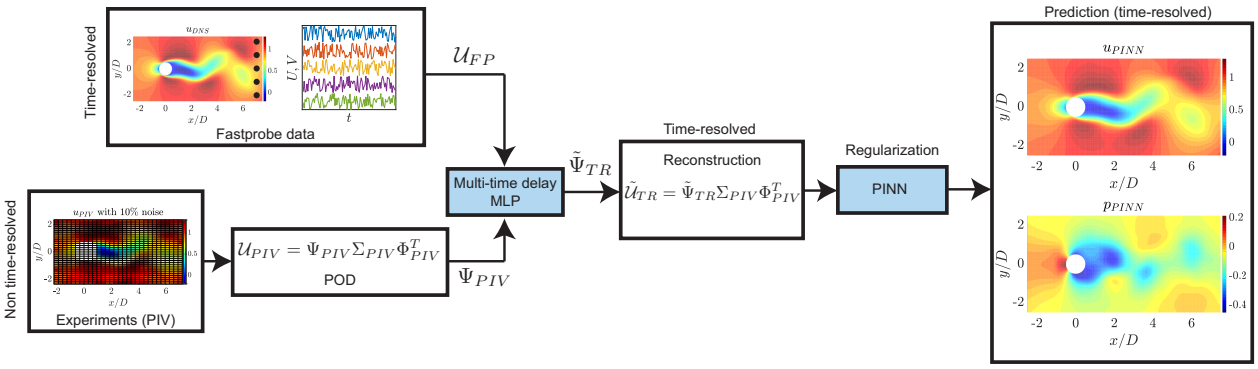

Figure 1. Sketch of the proposed architecture applied to the wake of a cylinder (for illustration purposes).

interval of 0.1 simulation units) of a fluid flow at $Re = 130$ (i.e. in the so-called chaotic regime) with x and y axis ranging from $[-5D, 15D]$ and $[-5D, 5D]$, respectively.

To compute time-resolved pointwise fast probe information, five virtual fast probes are located at the downstream edge of the domain ($x \approx 15D$), equally spaced along the vertical direction. This replicates a realistic configuration in which the intrusiveness of point probes should be minimized.

Synthetic PIV data are obtained by randomly seeding the velocity fields with particles and adding random noise up to 10%. The particle seeding density is set to 0.02 particles per pixel. The domain is discretized in 1024×512 pixels, thus corresponding to 51.2 pixels/D. A reasonable approximation of a PIV processing can be obtained by considering a moving average of the velocity sampled by the particles. The synthetic PIV interrogation window is set to 32×32 pixels, with a 50% overlap.

The point probes are recording velocity and pressure information at full time resolution, i.e. the data corresponding to the probe location is saved for the 20,000 DNS snapshots. Non-time resolved PIV snapshots are generated by considering a time separation of 200 DNS time steps. This interval corresponds to one flow throughtime, i.e. a fluid parcel entering the domain takes in average 200 snapshots to leave it. This originates from the fact that the convective dimensionless velocity corresponds to $1D$ per simulation unit, the length of the field of view is $20D$ and the interval between snapshots is 0.1 simulation units. A total of 100 PIV snapshots is thus generated. The synthetic PIV fields are analyzed with POD to extract the temporal modes Ψ_{PIV}. In order to reduce the dimensionality of the problem, a rank truncation is performed using the elbow method (see e.g. Marutho *et al.* (2018)). For this test case, the identified rank is $R = 12$. The main advantage of using POD as an encoder is the dimensionality reduction of the output, being it fixed at R coefficients corresponding to the most energetic modes.

The temporal mode estimator, i.e. the MLP, is fed by probe data corresponding to the 200 time instants past each considered PIV snapshot. As detailed in the Methodology section, in convection-dominated flows this correspond to a multi-time delay approach. The selected timespan correspond to one convective flow throughtime, as suggested by Discetti *et al.* (2018). The MLP architecture consists of a set of fully-connected dense layers of decreasing number of neurons. More specifically, the number of neurons decreases from an initial amount of $3 \times N_{probes} \times 200$, where $N_{probes} = 5$ corresponds to the number of fast probes installed in the domain and 200 matches the number of time instants between each PIV snapshot, to end up in a final dense layer with dimensions the number of temporal modes to estimate, i.e. $R = 12$. We opted for an Adam optimizer with an initial learning rate of 10^{-3} which is subsequently decreased parallel to the decrease of the training loss to allow for the optimal reconstruction based on empirical tests. This pyramidal structure allows for a fast and steady convergence given the significant reduction of the order of magnitude between the dimensionality of the input as compared to the output variables.

Once the temporal resolution of the probes has been embedded into the velocity fields with the MLP, we implement a PINN to enhance the accuracy by adding physical constraints. Our model for PINN consists on a fully-connected neural network with 10 hidden layers, each of them containing 100 nodes for each output variable, i.e. for a $2D$ domain, we have 300 nodes corresponding to the two components of velocity and pressure. The same optimizer as for the MLP is implemented here. We use for training of the PINN the outcome of the MLP on the test dataset. This is representative of the case of direct use of PINNs from probe data, i.e. once the flow estimator has been trained. Clearly this same process could have been carried out by considering the training data, although in this case higher accuracy is expected for in-sample snapshots than for time instants in between two PIV snapshots. In addition, to allow for a full reconstruction and extraction of the pressure field (note again that this is an additional feature, not available in the first place from PIV measurements), we use as boundary condition the pressure values given at the position of the probes, as if they were able to measure velocity and pressure simultaneously. We highlight that this boundary condition could have been exchanged by any other pressure measurement at any point within the domain.

RESULTS

Figure 2 shows the contour representation of the streamwise and crosswise components of the velocity field for an instantaneous snapshot with its corresponding pressure field. We include for the clarity of the comparison its associated PIV fields, even though this information is not included in the training dataset. The synthetic PIV field is in this case projected on the POD basis computed from the training dataset, and truncated to include the same number of modes R estimated by the MLP. The flow field estimated with the MLP and the regularized field after applying the PINN are also included. The pressure field of the MLP is obtained from integration of the pressure gradient computed with the momentum equation. A robust iterative procedure is carried out, following the implementation by Chen *et al.* (2022). Even though the MLP is already capable of achieving a good reconstruction of the velocity fields, significant discrepancies with respect to the refer-

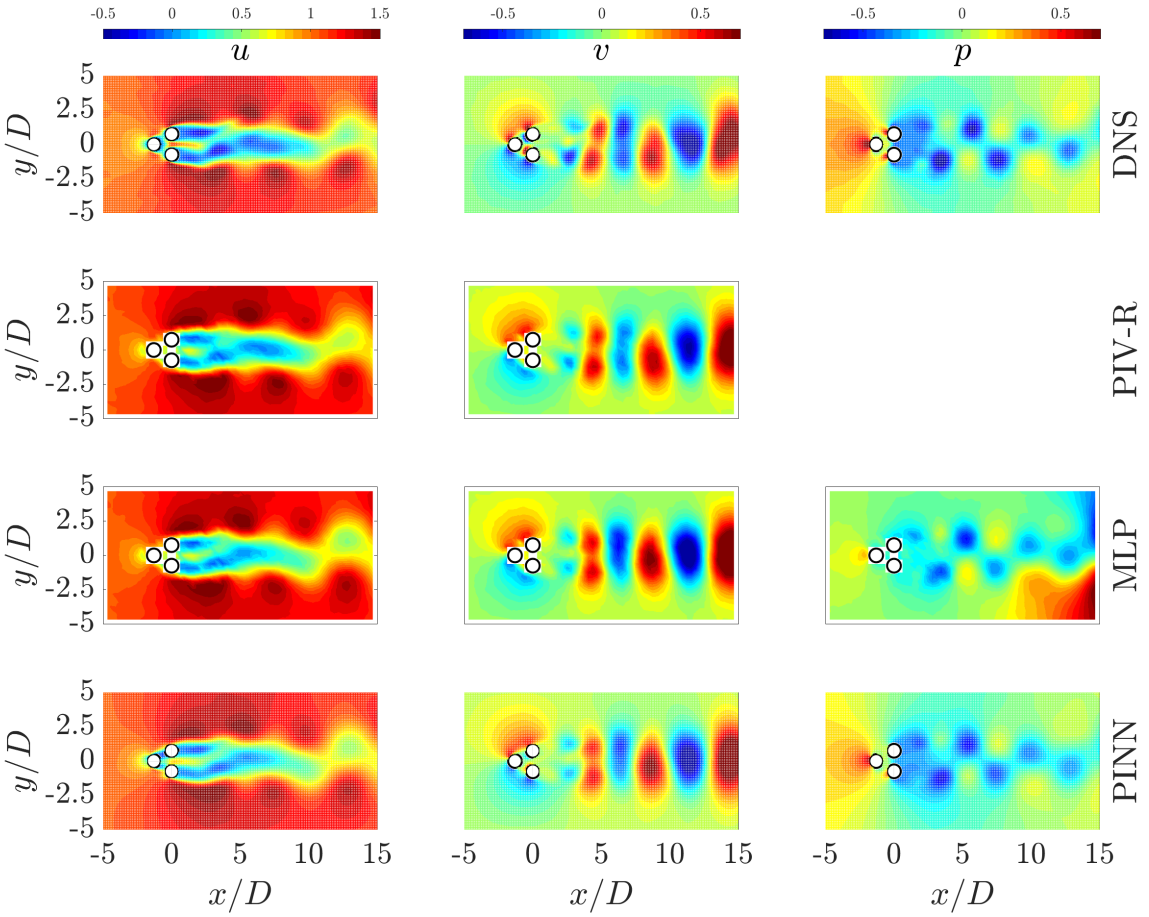

Figure 2. Reference and reconstructed fields after each stage of the proposed architecture on the test dataset: (first row) DNS reference velocity and pressure fields, (second row) reference PIV velocity field truncated to the best rank R (bear in mind that this row only exists for reference purposes and actually no PIV information is available once the model is trained), (third row) velocity field estimation after the MLP stage receiving fast probe measurements as input and reconstructing from the predicted temporal modes (note that pressure has been computed directly by solving NS equations and therefore, carry the errors from the estimation using MLP), and (fourth row) final regularized velocity and pressure fields after PINN.

ence DNS arise when computing pressure. The main sources of error arise from the finite spatial resolution of the PIV fields used for training (thus losing resolution in presence of large velocity gradients, especially in the near field), from truncation to rank R of the POD coefficients, and from temporal inconsistencies due to the finite accuracy of the estimation process. On the other hand, the PINN is able to achieve a more accurate pressure field estimation by imposing physical constraints in the reconstruction process.

The training error with respect to each of the contributions of the loss function is depicted in figure 3. The decrease of the total loss function indicates that the reconstructed fields from the MLP are being continuously regularized by enforcing compliance with the Navier-Stokes equations with the PINN. However, the figure shows that a plateau is reached after less than 100 epochs and more iterations would not result in an improvement in accuracy. In our experience, this plateau results from a compromise between the tendency of the Navier-Stokes regularization to homogenize the field whereas the estimated field data from the MLP and the pressure boundary condition push the network to comply with experimentally-accessible data. In the end, the overall error after this regularization is reduced significantly, as apparent in figure 2. A quantitative analysis is also reported in figure 4, where the spatially-averaged squared error is included for each snapshot of the testing dataset and the spatiotemporally-averaged squared error of the same dataset is presented in table 1. A clear reduction of the error with respect to the estimation with the MLP is observed. Most remarkably, PINNs are able to smooth out snapshots with large errors, which would have inevitably compromised the computation of temporal derivatives. This is a clear advantage, considering that the estimation of derived quantities often relies on the estimation of time derivatives from temporally-resolved data (see for instance pressure estimation (van Oudheusden, 2013)).

The PINN also provides additional information of the fluid flow, such as the pressure field. We observe in figure 2 that the pressure field is in qualitative excellent agreement with the pressure field of the original simulation, with most of the discrepancies occurring in the region between the cylinders and in the near field. Figure 4 shows indeed that a major improvement is achieved as compared to a direct calculation of the pressure distribution from the fields estimated with the MLP. It must be remarked, however, that the PINN implementation allows to impose directly compliance with measured pressure in more than one point (in this case, in the 5 probe locations). For the pressure derivation with the velocity fields estimated by the MLP, instead, we have simply imposed the value of the pressure in one node for the necessary boundary

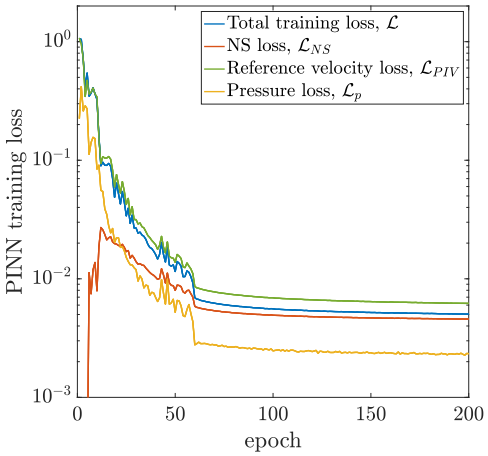

Figure 3. Loss function evolution during the different epochs of the PINN training process. Each contribution, namely, residuals of the Navier-Stokes equations and errors with respect to velocity and pressure reference data, are indicated. The continuous decrease of the loss function guarantees the convergence to a more accurate solution than the one resulting after the MLP.

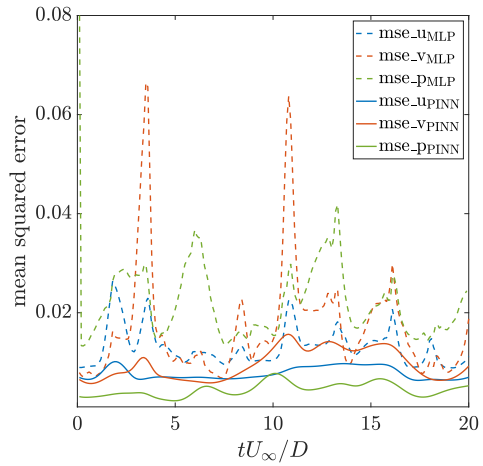

Figure 4. Mean square error (mse) of the velocity fields with respect to the DNS for the reconstruction after the MLP (dashed blue - u - and red - v - curves) and the regularization with PINNs (solid blue and red curves) over one through-time of the domain (corresponding to 200 snapshots at DNS time resolution). The green curves show the mse of the pressure with respect to DNS before (dashed, calculated directly from integration of the pressure gradient computed imposing the momentum equation with the velocity fields estimated via MLP) and after (solid) the regularization via PINNs.

condition for the integration of the pressure gradient.

The error distribution between MLP and PINN with respect to the DNS for an instantaneous snapshot is illustrated in figure 5, showing an excellent improvement of the accuracy of all the reconstructed fields.

CONCLUSIONS

We have proposed a solution to enable the estimation of flow fields from a small number of sensors in convection-dominated flows, and increase the accuracy by embedding physical constraints with a physics-informed neural network. The proposed architecture leverages the capability of a MLP to estimate temporal modes of POD and the output regularization enforced by the PINN. POD allows for the identification of a compact low-dimensional embedding of the most relevant features of the flow fields, thus permitting a fast and effective training with an MLP fed by probes. The input information is also enhanced using a multi-time delay approach, where temporal information collected by probes is enforced as spatial information on grounds of the validity of Taylor's hypothesis. This approach has been shown to work efficiently in convection-dominated flows with estimation through extended POD, and it is now enhanced with non-linear mapping through a MLP.

Table 1. Spatiotemporally-averaged squared error of the complete test dataset between the reference DNS and the estimation after the MLP and the final regularization via PINNs, correspondingly.

Feature	mse(DNS, MLP)	mse(DNS, PINN)
u	0.0174	0.0079
v	0.0227	0.0097
p	0.0432	0.0045

Additionally, the proposed concept allows to use PINNs for snapshot PIV data, i.e. in absence of temporal resolution, embedding in the process also equations including temporal derivatives (such as the momentum equations). This enables extracting pressure fields by including proper boundary conditions in the training process. The obtained results show a significant accuracy enhancement of the predicted velocity fields from MLP after the PINN regularization, and a reasonably good accuracy also for pressure field estimations. Pressure estimation could have also being achieved directly from estimated fields, although the accuracy of this process suffers due to truncation errors of the POD basis and inaccuracies of the mapping from probe to field data, among other error sources.

Our results show that PINNs with the complete set of Navier-Stokes equations can be used in conjunction with probe data and a MLP estimator to obtain velocity fields from probe data with a reasonably good accuracy. It must be remarked, however, that the validation test case is based on a simulation at relatively small Reynolds number of a wake flow. In this configuration, POD is able to provide a compact basis, thus easing the task of the estimation of the velocity fields. This concept is to be tested on flows with a richer spectrum of scales, in which the amount of modes to be reconstructed might pose a significant challenge.

ACKNOWLEDGMENTS

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No 949085) and by MCIN / AEI / 10.13039/501100011033 and the European Union 'NextGenerationEU/PRTR' as part of the grant FJC2020-044342-I. The authors warmly acknowledge N. Deng, B. Noack, M. Morzynski and L. Pastur for providing the code for the fluidic pinball simulation.

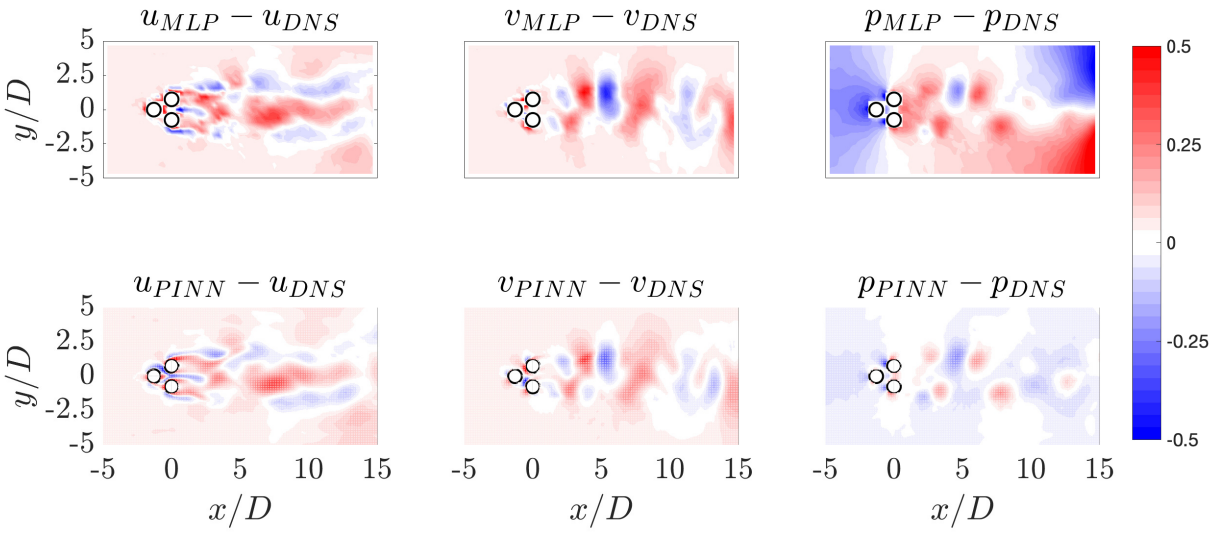

Figure 5. Error with respect to the DNS of each feature of the flow field after (first row) MLP reconstruction on the PIV grid and (second row) PINN regularization on a fine target grid (same as DNS). The improvement after the PINN is significant as compared to the output of the MLP, more specifically regarding the pressure field calculated by direct application of NS to the estimated velocity fields.

REFERENCES

Borée, J. 2003 Extended proper orthogonal decomposition: a tool to analyse correlated events in turbulent flows. *Exp. Fluids* **35**, 188–192.

Cai, S., Wang, Z., Wang, S., Perdikaris, P. & Karniadakis, G. E. 2021 Physics-informed neural networks for heat transfer problems. *J. Heat Transfer* **143(6)** (060801), 1–15.

Chen, J., Raiola, M. & Discetti, S. 2022 Pressure from data-driven-estimated velocity fields using snapshot PIV and fast probes. *Exp. Therm. Fluid Sci.* **136** (110647), 1–13.

Deng, N., Noack, B. R., Morzyński, M. & Pastur, L. R. 2020 Low-order model for successive bifurcations of the fluidic pinball. *J. Fluid Mech.* **884** (A37), 1–41.

Deng, Z., Chen, Y., Liu, Y. & Kim, K. C. 2019 Time-resolved turbulent velocity field reconstruction using a long short-term memory (LSTM)-based artificial intelligence framework. *Phys. Fluids* **31** (075108), 1–12.

Discetti, S., Bellani, G., Örlü, R., Serpieri, J., Sanmiguel Vila, C., Raiola, M., Zheng, X., Mascotelli, L., Talamelli, A. & Ianiro, A. 2019 Characterization of very-large-scale motions in high-Re pipe flows. *Exp. Therm. Fluid Sci.* **104**, 1–8.

Discetti, S., Raiola, M. & Ianiro, A. 2018 Estimation of time-resolved turbulent fields through correlation of non-time-resolved field measurements and time-resolved point measurements. *Exp. Therm. Fluid Sci.* **93**, 119–130.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. & Schmidhuber, J. 2017 LSTM: a search space odyssey. *IEEE Trans. Neural Networks Learn. Syst.* **28** (10), 2222–2232.

Han, J., Kim, D., Shin, H. & Kim, K. C. 2021 Enhanced data assimilation of 4D LPT with physics informed neural networks. In *14th International Symposium on Particle Image Velocimetry - ISPIV 2021*.

Hochreiter, S. & Schmidhuber, J. 1997 Long short-term memory. *Neural Comput.* **9** (8), 1735–1780.

Hosseini, Z., Martinuzzi, R. J. & Noack, B. R. 2015 Sensor-based estimation of the velocity in the wake of a low-aspect-ratio pyramid. *Exp. Fluids* **56** (13), 1–16.

Jin, X., Laima, S., Chen, W.-L. & Li, H. 2020 Time-resolved reconstruction of flow field around a circular cylinder by recurrent neural networks based on non-time-resolved particle image velocimetry measurements. *Exp. Fluids* **61** (114), 1–23.

Kalchbrenner, N., Danihelka, I. & Graves, A. 2015 Grid long-short term memory. *arXiv* .

Mao, Z., Jagtap, A. D. & Karniadakis, G. E. 2020 Physics-informed neural networks for high-speed flows. *Comput. Methods Appl. Mech. Eng.* **360** (112789), 1–26.

Marutho, D., Handaka, S. H., Wijaya, E. & Muljono 2018 The determination of cluster number at k-mean using elbow method and purity evaluation on headline news. In *2018 International Seminar on Application for Technology of Information and Communication*, pp. 533–538.

van Oudheusden, B. W. 2013 PIV-based pressure measurement. *Meas. Sci. Technol.* **24** (032001), 1–32.

Raissi, M., Perdikaris, P. & Karniadakis, G. E. 2019 Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J. Comput. Phys.* **378**, 686–707.

Raissi, M., Yazdani, A. & Karniadakis, G. E. 2020 Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations. *Science* **367** (6481), 1026–1030.

Sirovich, L. 1987 Turbulence and the dynamics of coherent structures. II. Symmetries and transformations. *Quart. Appl. Math.* **45** (3), 573–582.

Tinney, C. E., Ukeiley, L. S. & Glauser, M. N. 2008 Low-dimensional characteristics of a transonic jet. Part 2. Estimate and far-field prediction. *J. Fluid Mech.* **625**, 53–92.

Wang, H., Liu, Y. & Wang, S. 2022 Dense velocity reconstruction from particle image velocimetry/particle tracking velocimetry using a physics-informed neural network. *Phys. Fluids* **34** (017116), 1–15.

Yang, L., Meng, X. & Karniadakis, G. E. 2021 B-PINNs: bayesian physics-informed neural networks for forward and inverse PDE problems with noisy data. *J. Comput. Phys.* **425** (109913), 1–23.