Predictive Modeling of Fluid Flows Using Conditional Score-Based Diffusion Models

Wilfried Genuist[1,2][0009−0001−6518−9655], Éric Savin[1][0000−0002−3767−0281], Filippo Gatti[2][0000−0001−7174−4048], and Didier Clouteau[2][0000−0002−4953−9184]

[1] Université Paris-Saclay, ONERA, 91120 Palaiseau, France
{wilfried.genuist, eric.savin}@onera.fr
[2] Université Paris-Saclay, CentraleSupélec, ENS Paris-Saclay, CNRS, LMPS–Laboratoire de Mécanique Paris-Saclay, 91190, Gif-sur-Yvette, France
{wilfried.genuist, filippo.gatti, didier.clouteau}@centralesupelec.fr

Abstract. With the advent of increased computational power and advanced numerical methods, the simulation of turbulent fields has become central to many contemporary disciplines. Traditional solvers, while delivering high-quality predictions, still struggle in accurately providing fast estimations of flows due to the complexity and inherently chaotic nature of turbulence. As numerous machine learning-based solvers have emerged to address this problem (*e.g.* physics-informed neural networks or operator learning algorithms), capturing intricate physical phenomena remains a significant challenge. In the realm of generative modeling, diffusion models have established new benchmarks for solving similar problems. In this context, we propose a model that leverages the power of conditional score-based diffusion models for fluid flow prediction. We also integrate an energy constraint that rely on the statistical properties of the flow, further enhancing the temporal stability of the simulation. Our research, centered on a highly turbulent dataset, revealed the remarkable stability and reliability of simple generative diffusion models for turbulent field prediction.

Keywords: Diffusion models · PDEs · Fluid mechanics. · Stochastic differential equations · Score matching · Numerical simulations

1 Context and Objective

Computational fluid dynamics (CFD) is a field of study focused on simulating and analyzing fluid flow behavior, modelled using partial differential equations (PDEs). This branch of fluid mechanics underpins the development of aerodynamic vehicles and renewable energy technologies, like wind turbines and hydroelectric systems, driving innovation and sustainability. Without external forces, most CFD problems are framed within the context of compressible fluid dynamics, governed by the Navier-Stokes equations for the conservation of mass, momentum, and energy, namely:

$$\begin{cases} \partial_t \varrho + \boldsymbol{\nabla}_{\mathbf{r}} \cdot (\varrho \mathbf{u}) = 0\,, \\ \partial_t (\varrho \mathbf{u}) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot (\varrho \mathbf{u} \otimes \mathbf{u} - \mathbf{S}) = \mathbf{0}, \qquad \forall (\mathbf{r}, t) \in \Omega \times]0, +\infty[\\ \partial_t (\varrho e_t) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot ((\varrho e_t - \mathbf{S})\mathbf{u} + \mathbf{q}) = 0\,, \end{cases} \tag{1}$$

where $\mathbf{u}(\mathbf{r},t)$ is the fluid velocity field at position $\mathbf{r} \in \Omega \subset \mathbb{R}^d$ and time $t \geq 0$; $\varrho(\mathbf{r},t)$ is the density; $\mathbf{S} = -P\mathbf{I} + 2\mu \boldsymbol{\nabla}_{\mathbf{r}} \otimes_s \mathbf{u}$ is the stress tensor where $P(\mathbf{r},t)$ is the pressure field, μ is the dynamic viscosity, and $\mathbf{I}$ is the identity matrix; $\mathbf{q}(\mathbf{r},\mathbf{t})$ is the heat flux vector due to thermal conductivity; and $e_t(\mathbf{r},t)$ is the total energy per unit mass.

The training data $\mathcal{D} = \{\mathbf{x}_0^0, \mathbf{x}_0^1, \dots \mathbf{x}_0^S\}$ with $\mathbf{x}_0^s = \{\mathbf{u}(\mathbf{r}_r, t_s), P(\mathbf{r}_r, t_s); 1 \leq r \leq R\}$, $0 \leq s \leq S$, $t_0 = 0$, and $t_S = \mathcal{T}$ in [5] are constituted by $S = 1000$ snapshots with $R = 128 \times 64$ pixels of $\mathbf{u} = (u,v)$ and P for a compressible Karman vortex street past a cylinder in two dimensions ($d = 2$) at transonic regime with Reynolds number $\mathrm{Re} = 10^4$ and varying Mach numbers $M \in [0.53, 0.63] \cup [0.69, 0.90]$. Our objective is to predict $\hat{\mathbf{x}}_0^\tau := \{(u(\mathbf{r}_r,\tau), v(\mathbf{r}_r,\tau), P(\mathbf{r}_r,\tau)); 1 \leq r \leq R\}$ for any unobserved time $0 \leq \tau \leq \mathcal{T}$ starting from arbitrary initial conditions at t_0. However, solving Eq. (1) in the turbulent regime usually involves complex numerical methods (*e.g.* direct numerical simulation) which provide high-fidelity outputs but come with significant computational costs.

To enable faster inference, machine learning solvers have made remarkable advancements [2], with growing theoretical evidence supporting their efficiency [7]. While capturing long-term phenomena remains a challenge for turbulence prediction, Physics-Informed Neural Networks (PINNs) [8] are increasingly bridging the gap but remain limited in generalization capability. More recently, generative models, and in particular diffusion models [4,7,9], are rapidly expanding their application in physics and gradually outpace other state-of-the-art techniques.

Diffusion models aim to infer $\hat{\mathbf{x}}_0^\tau$ from p_0, the posterior distribution of the data set $\mathcal{D}$. Building upon the work of Kohl et al. [5], inference from p_0 is done by autoregression, meaning that continuous time prediction relies on successive predictions of the simulation states conditioned on l previous states. The maximization objective reads:

$$\hat{\mathbf{x}}_0^\tau = \arg\max_{\mathbf{x}_0} \; p_0\left(\mathbf{x}_0 | \mathbf{c}(\tau, l)\right) , \tag{2}$$

conditioning p_0 on $\mathbf{c}(\tau,l) = (\mathbf{x}_0^{\lfloor\tau\rfloor-1}, \mathbf{x}_0^{\lfloor\tau\rfloor-2}, \dots \mathbf{x}_0^{\lfloor\tau\rfloor-l})$, the l last steps of a simulation. Here $\lfloor\tau\rfloor = s$ stands for the index s such that $t_s \leq \tau < t_{s+1}$.

2 Conditional Flow Field Prediction

Score-based diffusion models [9] are generative diffusion models that extend the concept of Denoising Diffusion Probabilistic Models [4] (DDPMs) to continuous time domain. These models originate from the non-normalized estimation of the density function of energy-based models through score function (also known as Stein's score function), which is defined as the gradient of the log-probability distribution: $\mathbf{s}(\mathbf{x}) = \boldsymbol{\nabla}_{\mathbf{x}} \log p(\mathbf{x})$.

Let $\{\mathbf{x}_t;\, 0 \leq t < \infty\}$ be a random process defined on $\mathbb{R}^p$ and indexed on $t \in \mathbb{R}_+$. It is a diffusion process if it is the solution of the Itō stochastic differential equation (SDE) [6] describing the evolution of a sample $\mathbf{x}_0$ to its final state $\mathbf{x}_T$ according to:

$$\mathrm{d}\mathbf{x}_t = \boldsymbol{\mu}(\mathbf{x}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(\mathbf{x}_t, t)\mathrm{d}\mathbf{W}_t \, , \quad \mathbf{x}_0 \sim p_0 \, , \tag{3}$$

where $\boldsymbol{\mu}(\cdot,t) : \mathbb{R}^p \to \mathbb{R}^p$ is the drift coefficient, $\boldsymbol{\sigma}(\cdot,t) : \mathbb{R}^p \to \mathbb{R}^{p\times p}$ is the diffusion coefficient and $\{\mathbf{W}_t; t \geq 0\}$ is a p-dimensional Brownian motion (or Wiener process). We can reverse the time $t \to T - t$ in (3), leading to the reverse-time SDE:

$$\mathrm{d}\mathbf{x}_t = \bar{\boldsymbol{\mu}}(\mathbf{x}_t, t)\mathrm{d}t + \boldsymbol{\sigma}(\mathbf{x}_t, t)\mathrm{d}\overline{\mathbf{W}}_t \, , \quad \mathbf{x}_T \sim p_T \, , \tag{4}$$

where $\bar{\boldsymbol{\mu}}(\mathbf{x},t) = \boldsymbol{\mu}(\mathbf{x},t) - \boldsymbol{\nabla}_{\mathbf{x}} \cdot \boldsymbol{\Sigma}(\mathbf{x},t) - \boldsymbol{\Sigma}(\mathbf{x},t)\boldsymbol{\nabla}_{\mathbf{x}} \log p_t(\mathbf{x})$, $\boldsymbol{\Sigma} = \boldsymbol{\sigma}\boldsymbol{\sigma}^\intercal$, p_t is the marginal probability density function of $\mathbf{x}_t$, and $(\overline{\mathbf{W}}_t)_{t\geq 0}$ is a time reversed Brownian motion.

Analog to discrete time approach (DDPMs), solving the reverse SDE (4) starting from a sample $\mathbf{x}_T$ from the converging (prior) distribution p_T leads to the generation of a new sample from p_0. This includes estimating the intractable time-dependent score function $\mathbf{s}(\mathbf{x}, t) = \boldsymbol{\nabla}_{\mathbf{x}} \log p_t(\mathbf{x})$ which is approached by a denoiser $\mathbf{s}_{\boldsymbol{\theta}}(\mathbf{x}_t, t)$ by means of score matching [9,10] leveraging the known transition kernel $p_{t|0}(\mathbf{x}_t|\mathbf{x}_0)$ (which assumes linear and time-dependent only coefficients, *e.g.* $\boldsymbol{\mu}_t = \boldsymbol{\mu}(\cdot, t)$ and $\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}(\cdot, t)$). Since the posterior distribution is conditional (*e.g.* $p_0(\mathbf{x}_0|\mathbf{c}(\tau, l))$), a conditional denoising estimator $\mathbf{s}_{\boldsymbol{\theta}}(\mathbf{x}_t, t, \mathbf{c}(\tau, l))$ is used [1], optimizing the loss function (for $s \in \{l+1, \ldots, S\}$):

$$\mathcal{L}_{\mathrm{DSM}}(\boldsymbol{\theta}) = \mathbb{E}_{\substack{t \sim \mathcal{U}(0,T) \\ \mathbf{x}_0, \mathbf{c}(\tau, l) \sim p_0(\mathbf{x}_0, \mathbf{c}) \\ \mathbf{x}_t \sim p_{t|0}(\cdot|\mathbf{x}_0)}} \left\{ \lambda(t) \left\| \mathbf{s}_{\boldsymbol{\theta}}(\mathbf{x}_t, t, \mathbf{c}(\tau, l)) - \boldsymbol{\nabla}_{\mathbf{x}} \log p_{t|0}(\mathbf{x}_t|\mathbf{x}_0) \right\|_2^2 \right\} ,$$

where $\lambda(\cdot)$ is set to be equal to $\operatorname{Tr} \boldsymbol{\Sigma}_t$ [9] to obtain a likelihood weighting function (optimal training strategy). Each sample being conditioned by its previous states, we propose adding a loss penalty related to the energy of the system for better time consistency. We start by splitting the flow velocity into its mean part $\mathbf{U}$ and a randomly fluctuating part $\mathbf{u}'$ such that $\mathbf{u}'(\mathbf{r}, t) = \mathbf{u}(\mathbf{r}, t) - \mathbf{U}(\mathbf{r}, t)$ with zero mean $\mathbb{E}\{\mathbf{u}'(\mathbf{r}, t)\} = \mathbf{0}$ [3]. For any position $\{\mathbf{r}\}_{\mathbf{r} \in \mathcal{R}}$ ($\mathcal{R}$ being all available positions outside the obstacle) and fields $\mathbf{u}(\cdot)$ from the dataset $\mathcal{D}$, the proposed loss penalty reads:

$$\mathcal{L}_{\mathcal{E}}(\boldsymbol{\theta}) = \sum_{\mathbf{r}, \mathbf{r}' \in \mathcal{R}} \left\| \mathbf{R}_{\mathbf{u}}(\mathbf{r}, \mathbf{r}', \tau - t_l, \tau) - \mathbf{R}_{\hat{\mathbf{u}}}(\mathbf{r}, \mathbf{r}', \tau - t_l, \tau) \right\|_2 ,$$

where $\mathbf{R}_{\mathbf{u}}(\mathbf{r}, \mathbf{r}', t, t') = \mathbb{E}\{\mathbf{u}'(\mathbf{r}, t) \otimes \mathbf{u}'(\mathbf{r}', t')\}$ is the 2×2 autocorrelation matrix function of the velocity field fluctuations and $\hat{\mathbf{u}}$ is the reconstructed estimated field (from $\hat{\mathbf{x}}_0 \simeq \mathbf{x}_t + \boldsymbol{\Sigma}_t . \mathbf{s}_{\boldsymbol{\theta}}(\mathbf{x}_t, t, \mathbf{c})$ [10]). Using Jensen's inequality, the weighted minimization objective $\mathcal{L}_{\mathrm{DSM}}^w$ is equivalent to (for $\lambda_{\mathcal{E}} \geq 0$):

$$\mathcal{L}_{\mathrm{DSM}}^w(\boldsymbol{\theta}) = \mathcal{L}_{\mathrm{DSM}}(\boldsymbol{\theta}) + \lambda_{\mathcal{E}} \, \mathbb{E}_{\substack{\mathbf{r}, \mathbf{r}' \in \mathcal{R} \\ \mathbf{u} \in \mathcal{D}}} \left\| \mathbf{u}'(\mathbf{r}, \tau - t_l) \otimes \mathbf{u}'(\mathbf{r}', \tau) - \hat{\mathbf{u}}'(\mathbf{r}, \tau - t_l) \otimes \hat{\mathbf{u}}'(\mathbf{r}', \tau) \right\|_2 .$$

3 Results

A U-net architecture (along with transformers[3]) is implemented, trained on 4 Nvidia HGX A100 with k-fold cross-validation. The inference is made on an extrapolation problem on Mach number in the set $\{0.50, 0.51, 0.52\}$ with $\mathcal{T} = 60$ which is notoriously hard to predict as shock waves occur frequently. We refer to [4,9] for optimal choice of parameters of both the SDE and model architecture. The influence of l is not discussed here but optimal findings suggest $l = 1$.

As shown in Fig. 1, the fluid flow prediction remains consistently stable over time[4], even though some fading occurs towards the end. This phenomenon however, adversely affects the Mean Square Error (MSE), making it less efficient compared to [5]. However, for early predictions, the MSE is low ($< 10^{-3}$), which is competitive. Despite its simplicity, this straightforward architecture achieves very good results given the complexity of the problem. Notably, our energy loss function has played a key role in

[3] Backbone of the model: https://huggingface.co/blog/annotated-diffusion and https://github.com/yang-song/score_sde_pytorch

[4] Animated results are available on: https://centralesupelec.fr/fluid_prediction and https://centralesupelec.fr/absolute_error

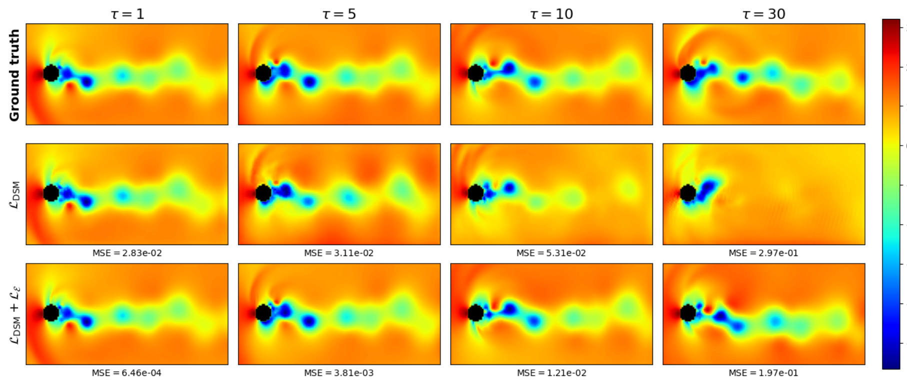

Fig. 1. Prediction of the pressure field at Mach 0.5.

ensuring the observed temporal stability and overall coherence of the intermediate predictions. While the results are satisfactory, there is still a need for further improvement in the conditioning process. Longer time horizon prediction are not feasible yet and the inference time is relatively long (around 1 hour). A plausible enhancement technique would be to perform diffusion in the latent space, which would substantially increase inference speed.

References

1. Batzolis, G., Stanczuk, J., Schönlieb, C.B., Etmann, C.: Conditional image generation with score-based diffusion models. arXiv:2111.13606 (2021)
2. Brunton, S.L., Noack, B.R., Koumoutsakos, P.: Machine learning for fluid mechanics. Annual Review of Fluid Mechanics **52**(1), 477–508 (2020)
3. Dryden, H.L.: A review of the statistical theory of turbulence. Quarterly of Applied Mathematics **1**(1), 7–42 (1943)
4. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems **33**, 6840–6851 (2020)
5. Kohl, G., Chen, L.W., Thuerey, N.: Benchmarking autoregressive conditional diffusion models for turbulent flow simulation. arXiv:2309.01745 (2023)
6. Oksendal, B.: Stochastic Differential Equations: an Introduction with Applications. Springer Science & Business Media (2013)
7. Qiu, J., Huang, J., Zhang, X., Lin, Z., Pan, M., Liu, Z., Miao, F.: Pi-fusion: Physics-informed diffusion model for learning fluid dynamics. arXiv:2406.03711 (2024)
8. Raissi, M., Perdikaris, P., Karniadakis, G.E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics **378**, 686–707 (2019)
9. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score-based generative modeling through stochastic differential equations. arXiv:2011.13456 (2020)
10. Vincent, P.: A connection between score matching and denoising autoencoders. Neural Computation **23**(7), 1661–1674 (2011)