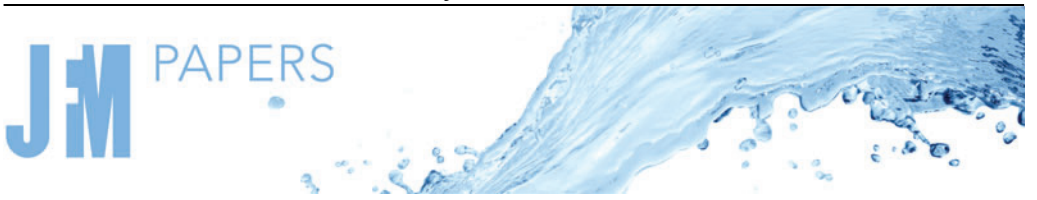
J. Fluid Mech. (2024), *vol*. 991, A5, doi:10.1017/jfm.2024.525



Jet mixing enhancement with Bayesian optimization, deep learning and persistent data topology

Yiqing Li[1], Bernd R. Noack[1,2,†], Tianyu Wang[1], Guy Y. Cornejo Maceda[1,†], Ethan Pickering[3], Tamir Shaqarin[4] and Artur Tyliszczak[5]

[1]Chair of Artificial Intelligence and Aerodynamics, School of Mechanical Engineering and Automation, Harbin Institute of Technology, 518055 Shenzhen, PR China

[2]Guangdong Provincial Key Laboratory of Intelligent Morphing Mechanisms and Adaptive Robotics, Harbin Institute of Technology, 518055 Shenzhen, PR China

[3]Independent Scholar

[4]Department of Mechanical Engineering, Tafila Technical University, 66110 Tafila, Jordan

[5]Faculty of Mechanical Engineering and Computer Science, Czestochowa University of Technology, 42-201 Czestochowa, Poland

(Received 3 November 2023; revised 25 March 2024; accepted 23 May 2024)

We optimize jet mixing using large eddy simulations (LES) at a Reynolds number of 3000. Key methodological enablers consist of Bayesian optimization, a surrogate model enhanced by deep learning and persistent data topology for physical interpretation. The mixing performance is characterized by an equivalent jet radius (R_{eq}) derived from the streamwise velocity in a plane located 8 diameters downstream. The optimization is performed in a 22-dimensional actuation space that comprises most known excitations. This search space parameterizes the distributed actuation imposed on the bulk flow and at the periphery of the nozzle in the streamwise and radial directions. The momentum flux measures the energy input of the actuation. The optimization quadruples the jet radius R_{eq} with a 7-armed blooming jet after around 570 evaluations. The control input requires 2 % momentum flux of the main flow, which is one order of magnitude lower than an *ad hoc* dual-mode excitation. Intriguingly, a pronounced suboptimum in the search space is associated with a double-helix jet, a new flow pattern. This jet pattern results in a mixing improvement comparable to the blooming jet. A state-of-the-art Bayesian optimization converges towards this double-helix solution. The learning is accelerated and converges to another better optimum by including a deep-learning-enhanced surrogate model trained along the optimization. Persistent data topology extracts the global and many local minima in the actuation space. These minima can be identified with flow patterns beneficial to the mixing.

† Email addresses for correspondence: bernd.noack@hit.edu.cn, yoslan@hit.edu.cn

Key words: mixing enhancement, machine learning, turbulence control

1. Introduction

Jet flows are ubiquitous in nature and technology and belong to a handful of configurations described in any fluid mechanics textbook. Jet mixing plays a pivotal role in many engineering applications, e.g. fuel injection in engines, combustor cooling, chemical mixing, printing and noise generation (Jordan & Colonius 2013), just to name a few. Hence, jet mixing optimization plays an important part in academic research and engineering applications.

Laminar jets are affected by the Kelvin–Helmholtz instability of the initial shear layer (Ball, Fellouah & Pollard 2012). The jet shear layer rolls up into pronounced vortex rings. Excitation at the nozzle exit provides authority over the vortex formation, e.g. allowing for the speed up of the vortex formation, to promote or mitigate vortex pairing, and to influence the far-field coherent structures. Vortex pairing in the streamwise direction promotes larger mixing regions observed as orderly 'vortical puffs' with axisymmetric excitation (Crow & Champagne 1971). More importantly, a significant increase in the spreading angle can be obtained by vortex splitting evolving along several branches (Lee & Reynolds 1985).

The actuation may promote axisymmetric, helical, dual-mode, flapping and bifurcating dynamics. In particular, acoustic excitation of the bulk affects the jet spreading via controlled vortex pairing (Crow & Champagne 1971; Hussain & Zaman 1980). Jet mixing is more effectively augmented with helical forcing (Mankbadi & Liu 1981; Corke & Kusek 1993). Bifurcating, trifurcating and blooming jets appear with a spreading angle up to 80° when axisymmetric and helical modes are combined (dual mode) with different frequency ratios (Lee & Reynolds 1985). The flapping mode is composed of counter-rotating helical modes, and the combination of axisymmetric and flapping modes is referred to as the bifurcating mode. Both the flapping and the bifurcating modes can produce bifurcating jets with impressive jet spreading (Parekh 1989; Danaila & Boersma 2000; da Silva & Métais 2002).

The world of multiple-mode actuation for mixing optimization holds considerable promise and is still to be explored. The radial excitation with three flapping modes, including 9 parameters, is optimized by evolution strategies (Koumoutsakos, Freund & Parekh 2001). Only one dominant flapping mode remains after 400 direct numerical simulations at $Re = 500$. In the sequel, the bifurcating mode using axial forcing is optimized for Re up to 1500 using the amplitudes and two Strouhal numbers as control parameters (Hilgers & Boersma 2001). In an adjoint-based optimization study at $Re = 2000$, radial forcing is found to be more effective than axial actuation in dual-mode forcing (Shaabani-Ardali, Sipp & Lesshafft 2020). In an experiment at $Re = 8000$, jet mixing is manipulated with periodic operation of six radial minijets. In 200 evaluations, Bayesian optimization minimizes a streamwise centreline velocity when tuning 12 parameters, the frequency, amplitudes and phase differences (Blanchard *et al.* 2021). The optimal mixing is facilitated by combining flapping and helical forcing, like machine learning control for the same configuration (Zhou *et al.* 2020). Moreover, the control performance also benefits from the deployment of more actuators and a richer actuation space. For example, an intelligent nozzle with eighteen electromagnetic flap actuators (Suzuki, Kasagi &

Suzuki 1999) and 8-channel localized arc filament plasma actuators (Utkin *et al.* 2006) have been developed for jet control.

In flow control, machine learning techniques have recently gained attention due to their successful applications (Duriez, Brunton & Noack 2017; Brunton, Noack & Koumoutsakos 2020). Examples are genetic programming and variants (Cornejo Maceda *et al.* 2021), reinforcement learning (Rabault *et al.* 2019; Guastoni *et al.* 2023; Nair & Goza 2023; Sonoda *et al.* 2023; Vignon *et al.* 2023*a*; Vignon, Rabault & Vinuesa 2023*b*; Xu & Zhang 2023) and Bayesian optimization (Blanchard *et al.* 2021). These methods encode the input–output relations in various forms without requiring prior knowledge. Function regression solvers like genetic programming and deep reinforcement learning can provide a large model capacity for exploration. However, deriving the optimal solution in finite time cannot be guaranteed. Alternatively, a predefined control law can be tuned to near optimal by parameter optimizers like Bayesian optimization (BO), genetic algorithm (GA) and particle swarm optimization (PSO), to name a few. Pino *et al.* (2023) compares genetic programming, deep reinforcement learning and BO in increasingly complex control problems. The authors highlight BO's potential to balance both sample efficiency and the performance of the final solution. With the recent advances in the design of the acquisition function (Blanchard & Sapsis 2021) and surrogate models (Pickering *et al.* 2022), BO is moving forward in conquering high-dimensional search spaces. This work leverages these advancements to optimize and understand high-dimensional jet forcing modes.

The present study builds on a jet mixing plant employing large eddy simulations (LES) and a rich streamwise and radial actuation space at the nozzle exit. This plant can reproduce virtually all previously considered actuated jet dynamics as elements of a high-dimensional search space. High-dimensional optimization constitutes a challenge that is tackled by a Bayesian optimizer enhanced by deep learning.

The paper is organized as follows. The configuration, actuation and metrics are defined in § 2. The optimizer and numerical solver are presented in § 3. We discuss the learning process and the optimized solutions in § 4. Finally, § 5 concludes the findings with outlook.

2. Set-up and problem definition

2.1. *Configuration and actuation*

The configuration is a jet flow exiting a circular nozzle of diameter D as shown in figure 1. The flow is described in a Cartesian coordinate system (x, y, z) where x represents the streamwise direction and the origin coincides with the centre of the nozzle. The computational domain starts from the exit and covers a rectangular region with size $12D \times 16D \times 12D$. The actuation $\boldsymbol{u}_b(r, \theta, t)$ is imposed with the mean streamwise velocity $u_m(r)$ as the inlet velocity profile $u(r, t)$

$$\boldsymbol{u}(r, t) = u_m(r)\boldsymbol{e}_x + \boldsymbol{u}_b(r, \theta, t), \tag{2.1}$$

where r measures the radial distance from the centreline, $\boldsymbol{e}_x$ is the unit vector in the x direction, and θ is the azimuthal angle. The mean streamwise component has a hyperbolic–tangent profile

$$u_m(r) = \frac{U_\mathrm{j} + u_c}{2} - \frac{U_\mathrm{j} - u_c}{2} \tanh\left(\frac{1}{4}\frac{R}{\delta_2}\left(\frac{r}{R} - \frac{R}{r}\right)\right), \tag{2.2}$$

where U_j is the jet centreline velocity, R is the radius of the nozzle, and $u_c = 0.03U_\mathrm{j}$ is the co-flow velocity to mimic a natural suction process and $\delta_2 = R/20$ is the momentum

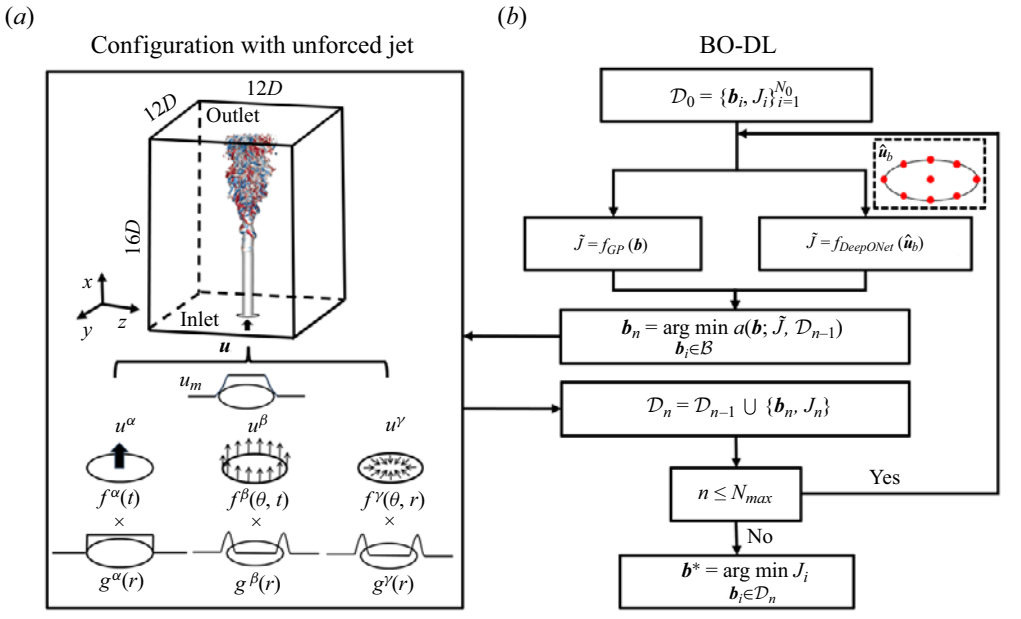


Figure 1. Problem set-up including the jet configuration with the designed actuation (*a*) and the deep-learning-enhanced BO (*b*). Here, $\boldsymbol{b}$ is the actuation parameter, $\hat{\boldsymbol{u}}_b$ is the actuation command at the 9 red points indicated in the dashed box.

boundary layer thickness of the initial shear layer. At the side boundaries, we impose that the vertical velocity equals u_c, and the remaining velocity components equal zero. The pressure at the side boundaries is computed from the Neumann condition $\boldsymbol{n} \cdot \boldsymbol{\nabla} p = 0$ with $\boldsymbol{n}$ as the vector normal to the boundary. At the outlet plane, the velocity is computed from a convective boundary condition $\partial \boldsymbol{u}/\partial t + \tilde{V}_C \partial \boldsymbol{u}/\partial n = 0$, where $\tilde{V}_C$ is the instantaneous convection velocity V_C limited to positive values: $\tilde{V}_C = \max(V_C, 0)$. Here, V_C is the velocity averaged over the outlet plane. The pressure at the outflow equals zero. Such a defined outflow boundary condition ensures stable simulations and has negligible impact on the turbulent flow structures leaving the computational domain (Tyliszczak & Geurts 2014; Tyliszczak 2018).

As introduced in § 1, the jet control techniques for mixing enhancement are usually designed according to the instability mode, described by their azimuthal wavenumber at order 0 (axisymmetric mode) or 1 (helical mode). The perturbation is either axial or radial. We combine both axial and radial perturbations and define the actuation $\boldsymbol{u}_b$ with a general expression of θ and t, without assumption on the forcing mode. Therefore, the term $\boldsymbol{u}_b$ includes an axisymmetric streamwise bulk forcing $u^\alpha(r, t)$, and a peripheral forcing with a streamwise component $u^\beta(r, \theta, t)$, and a radial $u^\gamma(r, \theta, t)$

$$\boldsymbol{u}_b = (u^\alpha + u^\beta)\boldsymbol{e}_x + u^\gamma \boldsymbol{e}_r. \tag{2.3}$$

The forcing components are the product of a perturbation $f^m(\theta, t)$ and a radial profile $g^m(r)$: $u^m(r, \theta, t) = f^m(\theta, t) g^m(r)$, $m = \alpha, \beta, \gamma$ with $g^\alpha(r) = 1$ for $r \le R$ and 0 for $r > R$, and $g^\beta(r) = g^\gamma(r) = \exp(-1000(R - r)^{2.5})$. The profiles of the three forcing components are depicted in figure 1. The perturbation terms $f^m(\theta, t)$ are defined as the sum and product

of space- and time-harmonic functions

$$f^{\alpha}(t) = \sum_{i=-L}^{L} \alpha_i \Theta_i(\omega_\alpha t), \tag{2.4}$$

$$f^{\beta}(\theta, t) = \sum_{i,j=-L,\ldots,L} \beta_{ij} \Theta_i(\theta) \Theta_j(\omega_\beta t), \tag{2.5}$$

$$f^{\gamma}(\theta, t) = \sum_{i,j=-L,\ldots,L} \gamma_{ij} \Theta_i(\theta) \Theta_j(\omega_\gamma t), \tag{2.6}$$

where α_i, β_{ij}, γ_{ij} and ω_α, ω_β, ω_γ are the actuation amplitudes and angular frequencies, respectively, and $\Theta_i(\phi)$ is the harmonic function basis defined as: $\Theta_i(\phi) = \sin(i\phi)$ for $i > 0$, $\Theta_i(\phi) = 1$ for $i = 0$, and $\Theta_i(\phi) = \cos(i\phi)$ for $i < 0$. The forcing ansatz can approximate any periodic function of θ and t as the expansion order increases. In this study, focus is placed on the first-order expansion ($L = 1$) of (2.4)–(2.6). Thus, the control law is parameterized by a 22-dimensional vector $\boldsymbol{b}$

$$\boldsymbol{b} = [St_\alpha, St_\beta, St_\gamma, \alpha_1, \{\beta_{ij}\}_{i,j=-1,0,-1}, \{\gamma_{ij}\}_{i,j=-1,0,-1}]^{\mathsf{T}} \in \mathcal{B}, \tag{2.7}$$

where the Strouhal numbers $St_m = \omega_m D/2\pi U_\mathrm{j}$, $m = \alpha, \beta, \gamma$. Note that α_0 is set to 0 as a constant bulk flow can be incorporated into the steady profile. In addition, $\alpha_{-1} = 0$ can be assumed by a translation in time. The range of St_m is set as [0.1, 1] to include the Strouhal number of the preferred mode at $St_p = 0.3$–0.64 (Crow & Champagne 1971; Gutmark & Ho 1983; Sadeghi & Pollard 2012), and the range of axisymmetric mode $St_\alpha \in [0.15, 0.8]$ where bifurcating and blooming jets are observed (Lee & Reynolds 1985; Parekh 1989; Tyliszczak 2018). The actuation amplitudes are limited to $-0.1 \leq \alpha_1, \beta_{ij}, \gamma_{ij} \leq 0.1$, lower than the 0.15 used by Danaila & Boersma (2000) and Gohil, Saha & Muralidhar (2015), Tyliszczak (2018) and 0.5 by Koumoutsakos *et al.* (2001).

This high-dimensional search space allows the actuation to emulate various forcing modes, such as axisymmetric, helical, flapping, bifurcating, dual mode and harmonic waves discussed in § 1. In short, the forcing can be axisymmetric, statistically axisymmetric and non-axisymmetric.

2.2. *Mixing and actuation performance*

The mixing process in turbulent jets is typically characterized by quantities such as the decay of centreline velocity and its fluctuations, or the entrainment (Nathan *et al.* 2006). However, these quantities only measure the local statistics and cannot reflect the mixing process of non-axisymmetric jets, such as bifurcating jets and asymmetric jets. For example, the velocity in these jets may locally drop to zero due to a jet-splitting phenomenon, rather than as a result of enhanced mixing. The entrainment is rather a measure of the amount of surrounding fluid entrained into the jet vicinity, without guaranteeing that it mixes with the jet. In asymmetric jets, the amount of fluid flowing towards the jet may characterize significant radial non-uniformity not revealed by the entrainment. Considering the non-axisymmetric forcing in search spaces, we define a new metric, the equivalent mixing radius R_{eq}, to estimate the spatial uniformity of the streamwise velocity. R_{eq} is defined as the normalized streamwise velocity variance

computed at a given y–z cross-section

$$R_{eq} = \sqrt{2}\sigma,\ \sigma^2 = \frac{\iint \varrho(y,z)[(y-y_c)^2 + (z-z_c)^2]\,\mathrm{d}y\,\mathrm{d}z}{\iint \varrho(y,z)\,\mathrm{d}y\,\mathrm{d}z}, \tag{2.8}$$

with $\varrho(y,z) = (\langle u(x = X_0, y, z)\rangle_t - u_c)/(U_j - u_c)$, $X_0 = 8D$, and (y_c, z_c) the jet centre, as an analogue to the centre of mass: $y_c = \iint y\varrho(y,z)\,\mathrm{d}y\,\mathrm{d}z/\iint \varrho(y,z)\,\mathrm{d}y\,\mathrm{d}z$ and $z_c = \iint z\varrho(y,z)\,\mathrm{d}y\,\mathrm{d}z/\iint \varrho(y,z)\,\mathrm{d}y\,\mathrm{d}z$. The $\sqrt{2}$ coefficient is chosen so that the equivalent mixing radius of a top flat jet flow of radius R is R.

The amplitude and mass flow rate have been adopted to evaluate the control input. Inspired by Parekh (1989), we define the momentum flux P of the actuation to estimate the energy input from a practical perspective. The momentum flux is time averaged and normalized by the jet axial momentum flux at the inlet

$$P = \frac{\left\langle \iint u_b^2\,\mathrm{d}A \right\rangle_t}{\pi R^2 U_j^2}. \tag{2.9}$$

3. Methodology

3.1. *Deep learning-enhanced Bayesian optimization*

The optimization problem to maximize the mixing as a response to the actuation input parameterized by $\boldsymbol{b}$ is formulated as

$$\boldsymbol{b}^* = \arg\min_{\boldsymbol{b}\in\mathcal{B}} J(\boldsymbol{b}), \tag{3.1}$$

where $\mathcal{B} = [0.1, 1]^3 \times [-0.1, 0.1]^{19} \subset \mathbb{R}^{22}$. Cost function J is defined as the inverse of the equivalent mixing radius and normalized by the unforced case, $J = R_{eq,0}/R_{eq}$. Better mixing with large R_{eq} is related to the decrease of J. The assumed optimization goal leads to the spatial uniformity of $\langle u(x = X_0, y, z)\rangle_t$.

To optimize this 22-dimensional search space, we employ techniques inspired by BO (Williams & Rasmussen 2006). Bayesian optimization has shown to be advantageous in optimizing expensive black-box functions by systematically reducing uncertainty in the black-box mapping and incorporating prior assumptions of the cost function (Shahriari *et al.* 2015). Through a sequential approach, BO identifies the next actuation to evaluate, or 'data point' to acquire, for the purpose of finding the optimum. This is generally achieved via a surrogate model trained on all the queried data and an acquisition function (Williams & Rasmussen 2006). A sketch of the method used in this study is shown in figure 1. The algorithm is initialized with the evaluation of a set $\mathcal{D}_0$ of N_0 actuation vectors in $\mathcal{B}$ generated by Latin hypercube sampling. Here, N_0 is equal to $N_D + 1$ with N_D the dimension of the search space $\mathcal{B}$. We recall that, for this study, $N_D = 22$. The set $\mathcal{D}_0$ includes all the evaluated parameter vectors and their cost $\{\boldsymbol{b}_i, J_i\}_{i=1}^{N_0}$. A surrogate model $\tilde{J}$ is trained on the available data to approximate the latent objective function J. After the initialization, the algorithm explores the search space $\mathcal{B}$ one new query at a time. At each iteration, BO determines the optimal actuation to implement next by minimizing an acquisition function $a(\boldsymbol{b})$. The acquisition function leverages the surrogate model $\tilde{J}$ and

available data $\mathcal{D}_{n-1}$ to guide the data selection in the search space. After each query, the data set is enriched by the new data point $\{\boldsymbol{b}_n, J_n\}$ into $\mathcal{D}_n$ to further refine the surrogate model. When the query budget is met, the algorithm ends with the best design vector $\boldsymbol{b}^*$ recorded during the optimization.

The two key elements in BO are the choice of the surrogate model and the sequential strategy (Blanchard & Sapsis 2021). We focus on the former for a better estimation of the high-dimensional flow control system. Gaussian processes (GP) serve as a successful surrogate model in moderate dimensions and can provide closed-form solutions with the posterior distribution. However, the computation of the posterior costs $O(n^3)$, where n is the number of observations due to the inverse of the covariance matrix. The number of evaluations required to effectively cover the domain grows exponentially with the dimensionality. This makes GP difficult to scale to large training sets for high-dimensional problems. Recently, the deep operator network (DeepONet) has shown small generalization error for systems where functions are acted upon by an operator (Lu *et al.* 2021). Different from GP, that parameterizes the input, DeepONet can map input functions, which are then transformed by an operator to an output function or scalar with improved accuracy. This means that DeepONet does not fall victim to the scaling difficulties of GP when training. Therefore, DeepONet is capable of learning from infinite-dimensional functions. Empirically, DeepONet's utility as an operator surrogate model for Bayesian-inspired experimental design has been shown to significantly outperform GP in several infinite-dimensional systems that exhibit extreme events in Pickering *et al.* (2022), ranging from stochastic pandemic spikes, to catastrophic structural failure, to rogue wave identification. Through a study of the Bayesian optimizer based on GP (BO) and DeepONet (BO-DeepONet) for the defined 22-dimensional problem (see § 4.2), we propose a new algorithm, deep learning-enhanced Bayesian optimization (BO-DL). By incorporating both GP and DeepONet as the surrogate model, BO-DL presents a better explorative capability and faster convergence. We alternate between DeepONet and GP every 10 iterations to query the next sample. The value of 10 is chosen empirically to balance the characteristics of the two models and combine their advantages. If the interval is too long, such as 100, the models will be more independent rather than interacting with each other. On the other hand, if the interval is too short, such as 1, the exploitation may be interrupted by uncertainty due to the exchange of models. In GP implementations, the parameter space of a stochastic process is used for both regression and searching. Instead, DeepONet performs regression in the functional space, leveraging the typically disregarded basis functions associated with the parameterization. Here, the function $\hat{\boldsymbol{u}}_b$ is designed as the actuation command at 9 points located at the jet exit, including the centreline and 8 equidistant points on the periphery.

The acquisition function employed is the likelihood-weighted lower confidence bound proposed by Blanchard & Sapsis (2021), with superiority in finding rare extreme behaviours

$$a(\boldsymbol{b}) = \mu(\boldsymbol{b}) - \kappa\sigma(\boldsymbol{b})w(\boldsymbol{b}), \quad w(\boldsymbol{b}) = \frac{p_{\boldsymbol{b}}(\boldsymbol{b})}{p_{\mu}(\mu(\boldsymbol{b}))}. \tag{3.2a,b}$$

Here, κ balances exploration (large κ) and exploitation (small κ), and is chosen as 1. The mean model $\mu(\boldsymbol{b})$ and standard variance model $\sigma(\boldsymbol{b})$ are estimated by the mean and variance over a 2-ensemble of trained DeepONet. For the case of GP, these can be calculated in closed form using standard expressions from GP regression (Williams & Rasmussen 2006). The likelihood ratio $w(\boldsymbol{b})$ measures relevance by weighting the uncertainty of the point (the input density $p_{\boldsymbol{b}}$) against its expected impact on the cost function (the output density p_{μ}).

3.2. *Governing equations and numerical solver*

We consider an incompressible flow described by the Navier–Stokes equations in the framework of LES

$$\frac{\partial \bar{u}_j}{\partial x_j} = 0, \tag{3.3}$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{1}{\rho}\frac{\partial \bar{p}}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j} + \frac{\partial \tau_{ij}^f}{\partial x_j}, \tag{3.4}$$

where u_i represents the velocity components, p denotes pressure and ρ is density. The overbar denotes spatially filtered variables, $\bar{f}(\boldsymbol{x}, t) = \int_\Omega f(\boldsymbol{x}', t)\mathcal{G}(\boldsymbol{x} - \boldsymbol{x}', \Delta)\,\mathrm{d}\boldsymbol{x}'$ and $\mathcal{G}$ is the filter function that fulfils the condition $\int_\Omega \mathcal{G}(\boldsymbol{x}, \Delta)\,\mathrm{d}\boldsymbol{x} = 1$. A local filter width equals the cube root of the computational cell volume, $\Delta = (\Delta x \Delta y \Delta z)^{1/3}$. The stress tensor includes the large-scale term τ_{ij} and the sub-grid term τ_{ij}^f defined as

$$\tau_{ij} = 2\nu S_{ij}, \quad \tau_{ij}^f = (\bar{u}_i\bar{u}_j - \overline{u_i u_j}), \tag{3.5a,b}$$

where ν is the kinematic viscosity and $S_{ij} = \frac{1}{2}(\partial \bar{u}_i/\partial x_j + \partial \bar{u}_j/\partial x_i)$ is the rate of strain tensor of the resolved velocity field. In this work, the sub-filter tensor is modelled by an eddy-viscosity model with $\tau_{ij}^f = 2\nu_t S_{ij} + \tau_{kk}^f \delta_{ij}/3$. The diagonal terms τ_{kk}^f are added to the pressure, resulting in the so-called modified pressure $\bar{P} = \bar{p} - \rho\tau_{kk}^f\delta_{ij}/3$. The Vreman subgrid-scale model is used for its low computational cost and very good accuracy in simulating jet flows (Wawrzak, Boguslawski & Tyliszczak 2015; Boguslawski, Wawrzak & Tyliszczak 2019).

The simulations are conducted with the in-house high-order LES solver SAILOR. The solution algorithm is based on the projection method for the pressure–velocity coupling for half-staggered meshes where the pressure nodes are shifted half a cell size from the velocity nodes (Tyliszczak 2014, 2015*a*). A predictor–corrector method (Adams–Bashforth/Adams–Moulton) is applied for the time integration. Derivative approximations and interpolation on staggered nodes are defined using sixth- and tenth-order compact difference formulas. The SAILOR solver has been used in jet studies with similar dynamic scales as the present work, such as jets undergoing laminar/turbulent transition (Boguslawski *et al.* 2019) and excited jets (Tyliszczak & Geurts 2014; Tyliszczak 2018). The applied high-order discretization schemes led to grid-independent results already with relatively coarse meshes.

4. Results

4.1. *Validation of the LES with a bifurcating jet*

The Reynolds number $Re = U_\mathrm{j}D/\nu$ is decided as 3000 for this study. This allows for the use of a relatively coarse computational mesh to obtain reliable and fast simulations as the database. Two meshes are employed in this study. A coarse mesh with $80 \times 160 \times 80$ nodes is used for the learning process and a refined mesh with $192 \times 336 \times 192$ nodes is used for the validation and flow analysis of selected cases. The mesh points are compacted in the axial direction towards the inlet using an exponential function and radially towards the jet axis by a tangent hyperbolic function. In the region $-1.2D < y, z < 1.2D$, the mesh spacing is almost uniform and equal to $\Delta y = \Delta z = 0.05D$ (46 nodes) on the

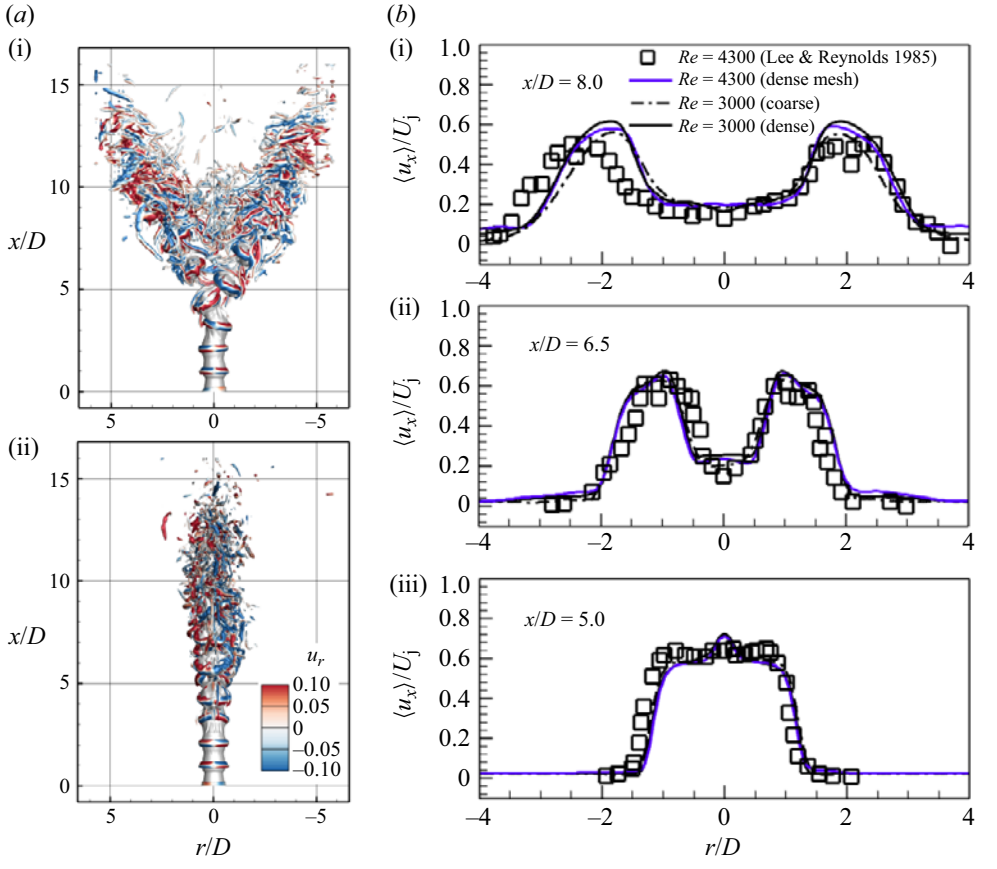


Figure 2. Validation of the LES solver on the bifurcating jet at $Re = 4300$ (Lee & Reynolds 1985). Instantaneous isosurfaces of Q-parameter ($Q = 0.5$) in the bifurcating (*a*-i) and bisecting planes (*a*-ii), coloured by the radial velocity u_r. Radial profiles of the time-averaged axial velocity in the bifurcating planes (*b*-i–iii). The square symbols represent the experimental results of Lee & Reynolds (1985), and the lines represent the simulation results obtained in this study.

coarse mesh and $\Delta y = \Delta z = 0.02D$ (115 nodes) on the dense one. In the axial direction, the sizes of the cells in the direct inlet vicinity are $\Delta x = 0.067D$ and $\Delta x = 0.032D$ for the coarse and dense meshes, respectively. The time step varies according to the Courant–Friedrichs–Lewy (CFL) condition, with the CFL number equal to 0.5. The jet impulsively injects into quiescent flow and becomes fully developed after $100D/U_j$ time units. The time-averaging procedure then starts and lasts for $500D/U_j$ time units for the statistics to converge. A single simulation on the coarse mesh takes 20 CPU-hours. The whole optimization process with 1000 converged simulations lasts around 21 days, using 40 CPUs of an AMD EPYC 7742 (2.25 GHz) processor. On the dense mesh, a single simulation run takes approximately 576 CPU-hours. The parallel computation is carried out with the Open MPI interface.

As presented in § 3.2, the numerical code employed has been well validated against experimental and numerical data for a series of studies of jet dynamics and control. Here, the code with the assumed perturbation design (2.4–2.6) is verified to obtain the well-documented flow pattern of an excited jet, the bifurcating jet at $Re = 4300$ (Lee & Reynolds 1985). The well-known jet is shown in figure 2, produced by the combined

axisymmetric and flapping excitation at the frequencies $0.4 < St_\alpha < 0.6$ and $St_\beta = St_\gamma = St_\alpha/2$. Based on the current control definition, the bifurcating jet is reproduced by

$$f^\alpha(t) = \alpha_1 \Theta_1(\omega_\alpha t), \tag{4.1}$$

that produces the axisymmetric excitation and

$$f^\beta(\theta, t) = \beta_{-1,1}\Theta_{-1}(\theta)\Theta_1(\omega_\beta t), \quad f^\gamma(\theta, t) = \gamma_{-1,1}\Theta_{-1}(\theta)\Theta_1(\omega_\gamma t), \tag{4.2a,b}$$

as the flapping mode, simulating the orbital motion of the nozzle tip in the experiment. We take $\alpha_1 = 0.17$, $St_\alpha = 0.5$ and assume $St_\beta = St_\gamma = St_\alpha/2$, $\beta^2_{-1,1} + \gamma^2_{-1,1} = \alpha^2_1$ with $\beta_{-1,1} = \alpha_1 \cos(20^\circ)$. This type of excitation is also used in the previous LES simulations of the bifurcating jet (da Silva & Métais 2002; Tyliszczak & Geurts 2014).

Figure 2(*b*-i–iii) shows the time-averaged axial velocity profiles along the radius in the bifurcating plane at the distance $x/D = 5.0,\ 6.5,\ 8.0$ from the inlet. These results were obtained for $Re = 4300$, as in Lee & Reynolds (1985), and for $Re = 3000$ assumed in the present study. The effect of the Reynolds number on the velocity profiles is small. We attribute such behaviour to a dominating role of the perturbation. The location and level of two peaks, which are associated with the split jet arms, are well predicted by the numerical solutions. The impact of the mesh density on the solution is also negligible, owing to the employed high-order numerical method. This also holds for the optimized jet, see figure 6 of § 4.4.

4.2. *Bayesian optimization with different surrogate models*

We first study the capability of the surrogate model, GP and DeepONet, to predict the cost function J as a response to the excitation input $\boldsymbol{b}$. Then, the Bayesian optimizers based on each of the two surrogate models (BO and BO-DeepONet) are tested on our plant. Finally, the performance of the proposed method BO-DL in § 3.1 is illustrated.

A k-fold cross-validation training ($k = 5$) of the GP and DeepONet model is performed over 1000 data points with 80/20 train/test split. The data are extracted randomly out of the database from realizations of Bayesian optimizers. Figure 3(*a*) shows the prediction $\tilde{J}(\boldsymbol{b})$ vs the truth $J(\boldsymbol{b})$ obtained. The distribution of points along the diagonal shows that DeepONet achieves a lower prediction error than GP. This is further explained by the correlation coefficient of $R = 0.89$ for DeepONet, and 0.71 for GP. The average error of the k tests is measured by the mean squared error (MSE). The MSE for GP model is 0.01, 1 % of the range of J value. The prediction of the DeepONet model is superior, with a MSE equal to 0.005. The learning process of the Bayesian optimizer with GP (BO) and DeepONet (BO-DeepONet) is shown in figure 3(*b*-i,ii). In figure 3(*b*-i), the learning curve of BO displays a plateau after the initial samples (triangles). After 160 samples, new optima are found and followed by continuous exploitation of the samples near the learning curve. The final solution is reached with $J = 0.274$ after 745 evaluations. When DeepONet is employed (figure 3*b*-ii), a better solution $J = 0.256$ is found quickly within 300 samples. This may be attributed to DeepONet's capability to generalize better for previously unseen data than GP, as the cross-validation indicates (Lu *et al.* 2021). After $m = 300$, the newly tested parameters cover the entire range of J, but no further improvement is observed in the learning curve. This suggests that the optimizer focuses on exploration of the search space rather than exploitation like BO. Based on the above observations, a joint surrogate model is proposed for this study to combine the advantages of GP in local exploitation and DeepONet in exploring new minima. The Bayesian optimizer based on this new model is described in § 3.1 and referred to as BO-DL. The learning process of BO-DL is given in

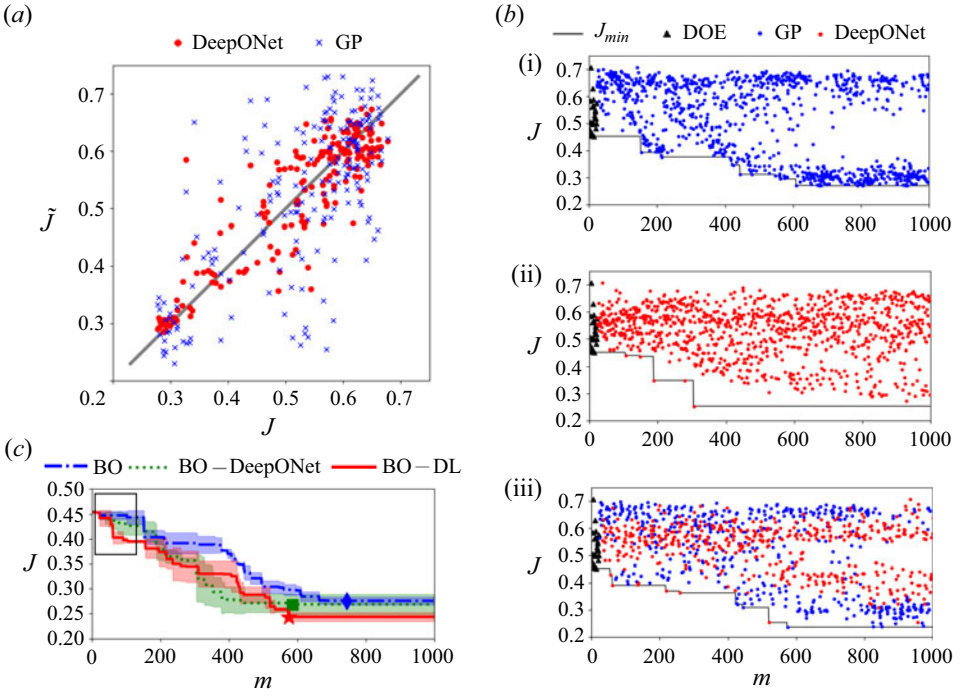


Figure 3. (*a*) Prediction of J by GP (crosses) and DeepONet (dots). Learning curves of J_{min} using BO (*b*-i), BO-DeepONet (*b*-ii) and BO-DL (*c*). (*d*) Average learning curves of Bayesian optimizers with/without deep leaning.

figure 3(*b*-iii) with the samples queried by GP and DeepONet. As indicated by the data points on the learning curve, the queries made by DeepONet (red dots) discovers a new minimum with significant reduction of J, and GP (blue dots) continues to descend. The best solution is obtained at $J = 0.237$ within 600 evaluations.

The average performance of the three Bayesian optimizers above is further studied. Each optimizer is employed for three realizations with a fixed budget of 1000 evaluations. Figure 3(*c*) reports the average value of the current optimum J_{min} from each optimizer with the standard deviation (shaded region) of three runs. The learning curve starts from $J_{min} = 0.45$, the lowest cost value after initialization of 23 samples, including the unforced case and the other 22 controlled cases from Latin hypercube sampling in the search spaces. The unforced case ($J = 1$) is omitted for better visibility of the data. The maximum cost of the controlled flow is around 0.7. With around 750, 580 and 570 queries, the average lowest costs J_{min} achieved by BO, BO-DeepONet and BO-DL are $J = 0.274$ (diamond), $J = 0.268$ (square) and $J = 0.237$ (star), respectively. On average, BO-DeepONet shows the fastest learning speed (dotted line) but with the largest variation. This is owing to DeepONet's capability of predicting the potential minima with a small generalization error. Although the descent of BO is the slowest, the optimal results of the three realizations are consistent. This indicates that GP provides better interpolation around the minima than DeepONet due to its deterministic nature. By combining GP and DeepONet, BO-DL not only demonstrates a comparable learning speed to BO-DeepONet but also inherits the small variance of the final solution from BO. Finally, among the three optimizers, BO-DL derives the best solution. In addition, the warm-up phase during queries 0 to 100 appears to be significantly shortened, denoted by the rectangle in figure 3(*c*).

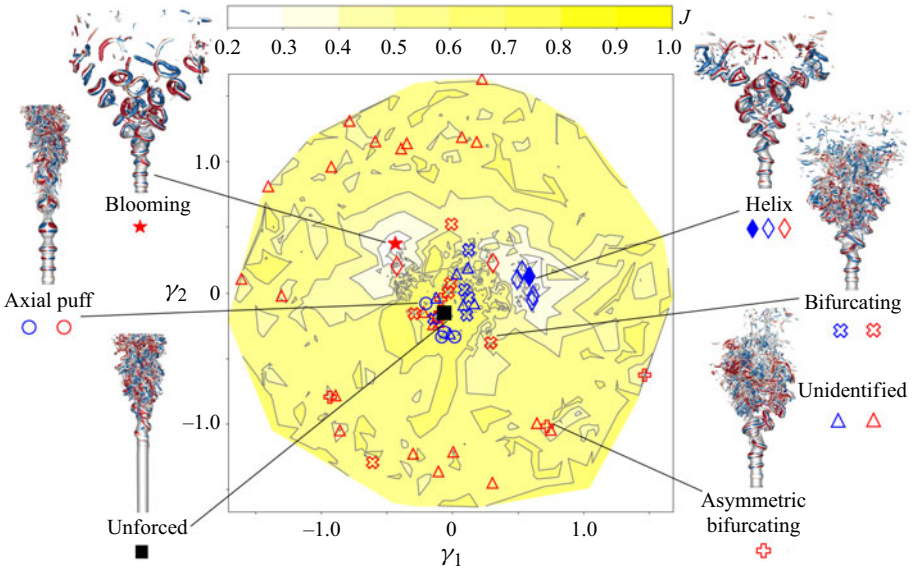


Figure 4. Learning process of BO and BO-DL on the proximity map. The unforced case (filled square), the local minima (unfilled symbols) and the final solutions explored by BO (blue-filled diamond) and BO-DL (red-filled star) are highlighted, with related jet patterns.

The computational cost of the BO loop is also noteworthy. With BO, the computation of the posterior costs is $O(N^3)$, where N is the number of observations (Williams & Rasmussen 2006). This makes the algorithm quite slow, even after only a few hundred observations. The experience of this study shows that the computation of BO increases from 10 CPU-seconds to 600 CPU-seconds after 1000 iterations on an AMD EPYC 7742 (2.25 GHz) processor. The BO-DeepONet procedure scales much more favourably. Initially, the first iteration takes 120 CPU-seconds, increasing only to 180 CPU-seconds after 1000 iterations. The combination of the two models in BO-DL compromises the cost to an average level.

For the high-dimensional physical problem, we show that BO can benefit not only from more accurate surrogate models but also from combining the advantages of parametric and non-parametric predictors. The proposed BO-DL holds a fast convergence and efficient exploration with a GP-DeepONet surrogate model. Compared with GP, the proposed surrogate model can provide more accurate predictions by leveraging the hidden functional input with DeepONet and scales better as both data size and dimensionality increase. In addition, a comparison between the Bayesian methods and bio-inspired approaches is given in the Appendix. It is shown that the optimizer with a surrogate model employed, particularly a deep-learning model, shows more advantage in the current problem.

4.3. *Exploration and characterization of the search space*

In this section, we explore the learning processes of the BO and BO-DL in the 22-dimensional space with persistent data topology (Wang, Cornejo Maceda & Noack 2023*a*; Wang *et al.* 2023*b*). This data analysis identifies the cost function minima and their depth, i.e. their persistence to noise, and was inspired by Edelsbrunner & Harer (2008). The analysis includes the identification of local minima in the high-dimensional actuation space, a dimension reduction to a two-dimensional proximity map and corresponding data visualization, as shown in figure 4. The 22-dimensional data obtained by both BO and

BO-DL are projected on a two-dimensional proximity map by classical multidimensional scaling. The feature coordinates γ_{ij} are chosen to optimally preserve the dissimilarity between control parameters defined by the Euclidean distance $D_{ij} = |\boldsymbol{b}_i - \boldsymbol{b}_j|$. The map features two large basins of attraction with low values of J, as well as small basins distributed around the border. A point $\boldsymbol{b}_0$ is supposed as a local minimum $\boldsymbol{m}$, if there exists a neighbourhood $\mathcal{B}$ of $\boldsymbol{b}_0$ that satisfies $J(\boldsymbol{b}_0) \leq \min_{\boldsymbol{b}\in\mathcal{B}} J(\boldsymbol{b})$. Here, $\mathcal{B}$ is an open set which should include the K nearest neighbours of $\boldsymbol{b}_0$ measured by Euclidean distance, $K \geq N_D + 1$. Note that the local minima are assumed based on the obtained discrete data and may change with additional data. A total of 57 local minima are extracted from the data, with 36 found by BO-DL and 21 by BO. In the proximity map, the unforced case is represented by a black square where both algorithms begin. The other symbols denote the derived minima $\boldsymbol{m}$ found by BO (blue) and BO-DL (red). The final BO and BO-DL solutions highlighted by the filled diamond ($J = 0.27$) and the filled star ($J = 0.24$) are located in the large basins of attractions. Most of the minima queried by BO are located in the centre of the map, whereas BO-DL also explores outward regions. Forced by the control commands corresponding to these minima, different jet patterns are observed, corroborated with the control modes. The axial puffs (circles) are close to the unforced case. The bifurcating type (cross) distributes widely in the cost range. The lower the J value is, the closer to the helix (filled diamond) basin. The jets bifurcating to one side are away from the centre, surrounded by the other unidentified patterns (triangles). Helix (diamond) and blooming (star) jets feature the most substantial performance, but the latter is only detected by BO-DL. Among the 20 minima explored by BO, the identified patterns include 6 helix, 5 flapping and 4 axial puffs. Among the 37 minima explored by BO-DL, the identified patterns include 6 flapping, 5 asymmetric flapping, 2 helix, 1 blooming and 1 axial puff. In addition, most (22) of the 27 unidentified patterns are detected by BO-DL.

The proposed BO-DL explores not only more minima than BO but also more diverse flow patterns beneficial to the mixing. This is probably owing to DeepONet's capability to extrapolate the mapping from the high-dimensional actuation to the mixing response more accurately. Two solutions with large basins of attractions in the search space are revealed – the optimal solution with a 7-armed blooming jet generated, and the suboptimal with a double-helix shape.

4.4. *Discussion of the optimized solutions*

Here, we include three solutions for the discussion: an *ad hoc* forcing with the best mixing in Tyliszczak (2018), BO optimized solution and the optimal solution of BO-DL. The forcing command, the instantaneous snapshots and the mean flow fields are presented in figure 5. The forcing commands are expressed by the operators in an order of constant, spatial periodic, temporal periodic and travelling waves in figures 5(*a*-i), 5(*b*-i) and 5(*c*-i). The axial excitation combining axisymmetric and helical modes has been widely employed to study the bifurcating and blooming jets since Lee & Reynolds (1985). A parametric study of the blooming jets with this type of excitation was performed in Tyliszczak (2018) under the same Reynolds number as this study. Among various multi-armed jets, the one with 11 arms led to the best mixing performance. The excitation was imposed on the axial velocity and combined the axisymmetric mode with the Strouhal number $St_a = 0.45$ and the helical mode with $St_h = 0.164$ at the same amplitude, 15 % of the bulk jet velocity (figure 5*a*-i). The BO solution contains mainly the axisymmetric mode at an amplitude of 8 % with $St_\alpha = 0.497$ for the bulk, a helical mode at an amplitude of 7 % and a flapping mode at an amplitude of 4 % with $St_\gamma = 0.232$ for radial components in the periphery.

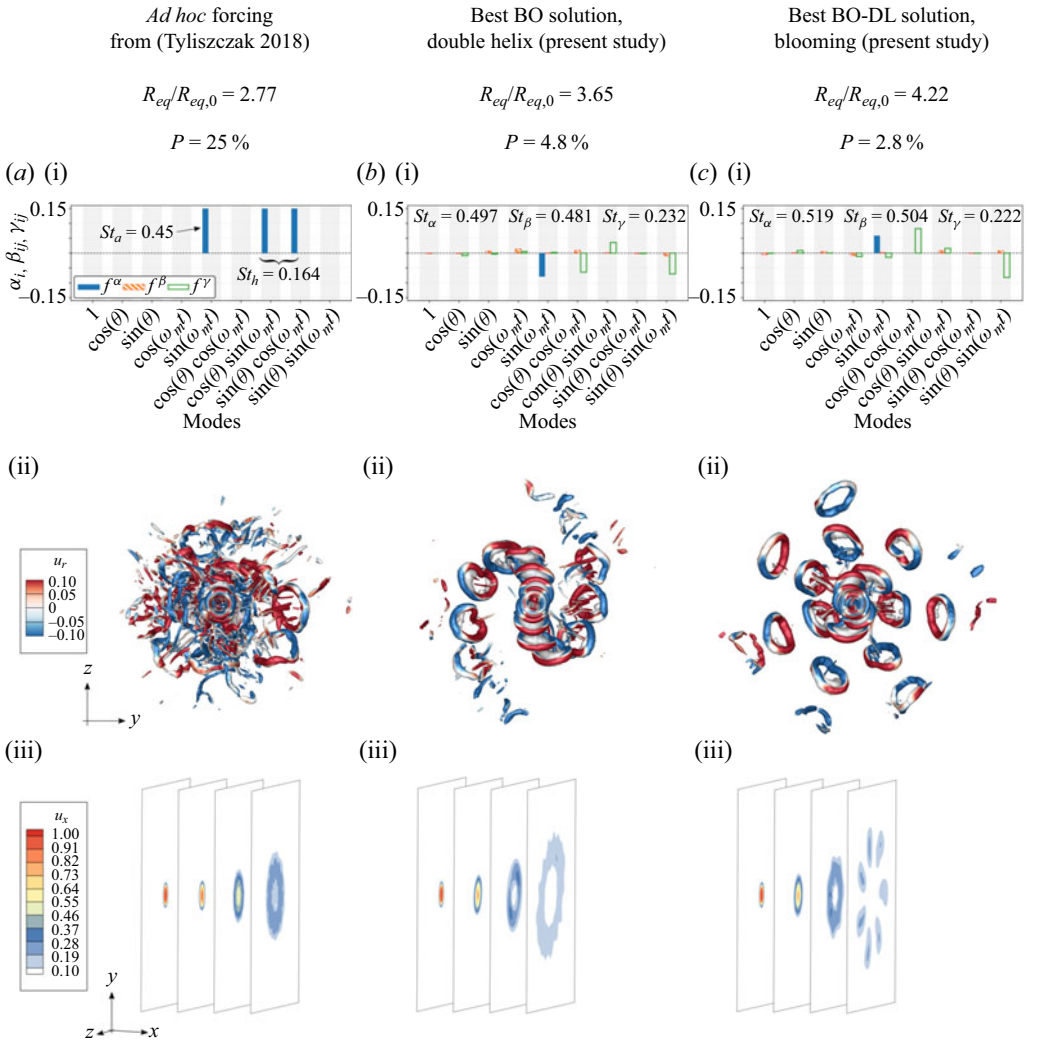


Figure 5. Comparison between three mixing jet solutions: (*a*) the *ad hoc* solution with the best mixing in Tyliszczak (2018), (*b*) the best solution learned with BO and (*c*) the best solution learned with BO-DL. Top – forcing modes with the associated mixing and actuation metric. Middle – (contour online) bottom view of instantaneous isosurfaces of $Q = 0.5$ coloured by the radial velocity u_r. Bottom – (contour online) contour plots of the streamwise velocity on cross-sectional planes at $x = 2D$, $4D$, $6D$, $8D$.

After removing the expansions with negligible amplitudes, less than 1 % of the bulk jet, the control law approximately reads

$$\left.\begin{aligned} f^\alpha(t) &= 0.08\sin(2\pi \times 0.497 U_j t/D), \\ f^\beta(\theta, t) &\approx 0, \\ f^\gamma(\theta, t) &\approx 0.04\cos(\theta)\sin(2\pi \times 0.232 U_j t/D) \\ &\quad -0.07\cos(\theta - 2\pi \times 0.232 U_j t/D). \end{aligned}\right\} \tag{4.3}$$

Because the removed terms hold an amplitude lower than the turbulence intensity at the jet outlet, the approximation hardly changes the flow patterns, with the relative cost difference being less than 1 %. The BO-DL solution contains mainly the axisymmetric mode at an amplitude of 6 % with $St_\alpha = 0.519$ for the bulk, a helical mode at an amplitude of 8 % with

$St_\gamma = 0.223$ for radial components in the periphery. The simplified control law reads

$$\left.\begin{array}{c} f^\alpha(t) = 0.06\sin(0.519t), \\ f^\beta(\theta, t) \approx 0, \\ f^\gamma(\theta, t) \approx 0.08\cos(\theta + 2\pi \times 0.223 U_j t/D). \end{array}\right\} \tag{4.4}$$

Two significant factors to be noted are the axisymmetric forcing Strouhal number St_α, and the frequency ratio between the axial and helical modes, $\alpha = St_\alpha/St_\gamma$. For both BO and BO-DL solutions, the axisymmetric forcing Strouhal number falls into the range $0.4 \lesssim St_\alpha \lesssim 0.6$ to observe bifurcating and blooming jets, and coincides with around $St_\alpha = 0.5$ where the peak spreading occurs (Lee & Reynolds 1985; Gohil *et al.* 2015; Shaabani-Ardali *et al.* 2020). The BO-DL actuation takes a frequency ratio of 2.34, which very well agrees with a theoretically derived value $\alpha = 7/3$ of Tyliszczak (2015*b*) and Gohil *et al.* (2015). Interestingly, the ratio of the BO solution which produces a helix jet ($\alpha = 2.14$) also falls into this range. Moreover, different from the *ad hoc* excitation using only the axial forcing, the radial component in the periphery plays an important role in solutions optimized by both BO and BO-DL. Shaabani-Ardali *et al.* (2020) also concludes radial forcing is the dominant component of helical modes to maximize the spreading angle of a bifurcating jet. We extend the importance of radial forcing to the jet spreading globally. From an estimate of the momentum flux, the solutions in this study take only 2.8 % (BO-DL) and 4.8 % (BO) of the main jet, one order lower than the *ad hoc* excitation (25 %). One reason is the low amplitudes, and another is the forcing applied into the local boundary region (see § 2.1) rather than the whole jet, which leads to a more efficient control. This represents the physical reality of small actuators installed on the wall of the inlet nozzle, like flap arrays in Suzuki *et al.* (1999), only affecting the boundary layers.

The flow structures are presented by the bottom view of the instantaneous Q-parameter isosurfaces (figures 5*a*-ii, 5*b*-ii, and 5*c*-ii). The arms of the *ad hoc* blooming jet are not explicitly observed due to the interaction between the closely aligned vortex rings. A double-helix jet is formulated by the BO solution. The jet bifurcates into two branches, which rotate with a specific frequency (see figure 8) and then experiences continuous bifurcation along the azimuth until the vortex rings break. This type of jet has not been reported in the literature so far. We reserve it for future investigation. The BO-DL optimized jet produces a 7-armed blooming jet, with the vortex rings eventually propagating along 7 different trajectories. The contour slices of the time-averaged streamwise velocity also confirm the spreading observed from the vortex rings. The 11 branches generated by the *ad hoc* forcing can be traced to $x = 8D$. The BO-optimized jet shows a more continuous distribution along the circumference due to the azimuthal bifurcation of two helix-shaped arms. The blooming jet is the earliest and furthest spreading. This leads to the largest effective mixing radius, $4.22R_{eq,0}$ at $x = 8D$, followed by BO optimized jet with $R_{eq} = 3.65R_{eq,0}$ and the *ad hoc* forced jet with $R_{eq} = 2.77R_{eq,0}$.

Figure 6 shows the axial profiles of the time-averaged centreline velocity and its fluctuation for the unforced and forced jets. The results obtained on the coarse and dense meshes agree well, except for slight discrepancies in the region $5D < x < 8D$. Compared with the unforced case in figure 6(*a*), the length of the potential core shortens significantly from $7.5D$ to $2D$ for both helical and blooming jets. Beyond the potential core, the velocity drops steeply and even reaches small negative values in the blooming jet. As with a similar behaviour observed by Tyliszczak (2018), this is the effect of jet splitting resulting in a local pressure drop. As a result, the reversal flow amplifies the local turbulence intensity to the peak at $x \approx 5D$, as indicated in figure 6(*b*). The initial fluctuation level in both jets corresponds to the imposed forcing amplitudes of the bulk forcing term. A small decrease

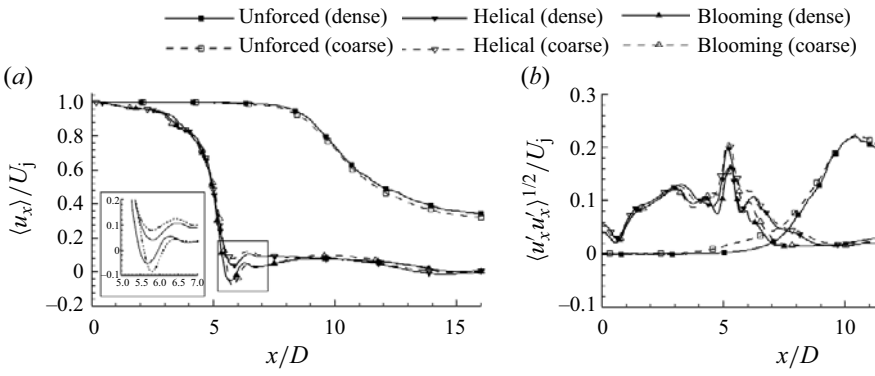


Figure 6. Mean profiles (*a*) and fluctuations (*b*) of the axial velocity for the unforced flow, the helical (BO solution) and the blooming jet (BO-DL solution). The solid and dashed lines denote the results calculated separately by the coarse and dense meshes.

at $x < 1D$ is caused by the lack of energy in a range of low wavenumbers (Kempf, Klein & Janicka 2005). For the controlled jets, the waves of the fluctuation profiles around the peak are attributed to the interactions between the Kelvin–Helmholtz instability and the forcing disturbance. Further downstream, the fluctuations drop nearly to zero along with a low mean velocity. The fluctuation profile for the unforced jet shows a very low turbulence level until $x \approx 7D$, and then slowly increases to the maximum around $x = 10D$.

Figure 7 shows the amplitude spectra of the centreline velocity at four locations along the axis, $x = 0D$, $2D$, $4D$ and $6D$. These results are presented vs the Strouhal number $St_D = \omega D/2\pi U_j$. The spectrum of the unforced jet is nearly flat at the inlet as the imposed turbulent signal does not contain any characteristic frequency. The high-frequency components ($St_D > 1$) are dampened downstream, and a broadband peak emerges around $St_D = 0.52$. This falls within the range of the preferred mode frequency $St_p = 0.3 - 0.64$ (Crow & Champagne 1971; Gutmark & Ho 1983; Sadeghi & Pollard 2012). Note that the optimal St_α predicted by BO-DL for the blooming jet perfectly matches the current preferred mode $St_p = 0.52$. This finding is consistent with previous studies (Tyliszczak & Geurts 2014; Gohil *et al.* 2015; Tyliszczak 2018) which concluded that the jet splitting phenomenon is most pronounced when St_α is equal to St_p. The initial spectra of the helix and blooming jets characterize a distinct peak at St_α. The peaks related to the helical forcing St_γ can be observed from $x = 2D$. The high-frequency harmonics also appear due to the interactions between generated toroidal vortices. In the case of the helical jet, the peak at $St_D \approx 0.032$ is also noteworthy. We find that this frequency coincides with the azimuthal motion of the helical arms, with a period T_r equal to $62.5D/U_j$. Figure 8 shows the snapshots of the helical jet, depicting the positions of its arms during the period of $31.25D/U_j$, which corresponds to the detected $St_D \approx 0.032$. The relationship between the frequency of the rotation and the one associated with forcing terms is left for future study.

5. Conclusions and outlook

We perform a global optimization of the jet control modes, parameterized in a 22-dimensional search space. The forcing includes axial and radial components that are defined to approximate a general periodic function of time and azimuthal angle. The design space allows the actuation to emulate various forcing modes that have been studied. This high-dimensional problem for jet mixing improvement is tackled by BO. We advance

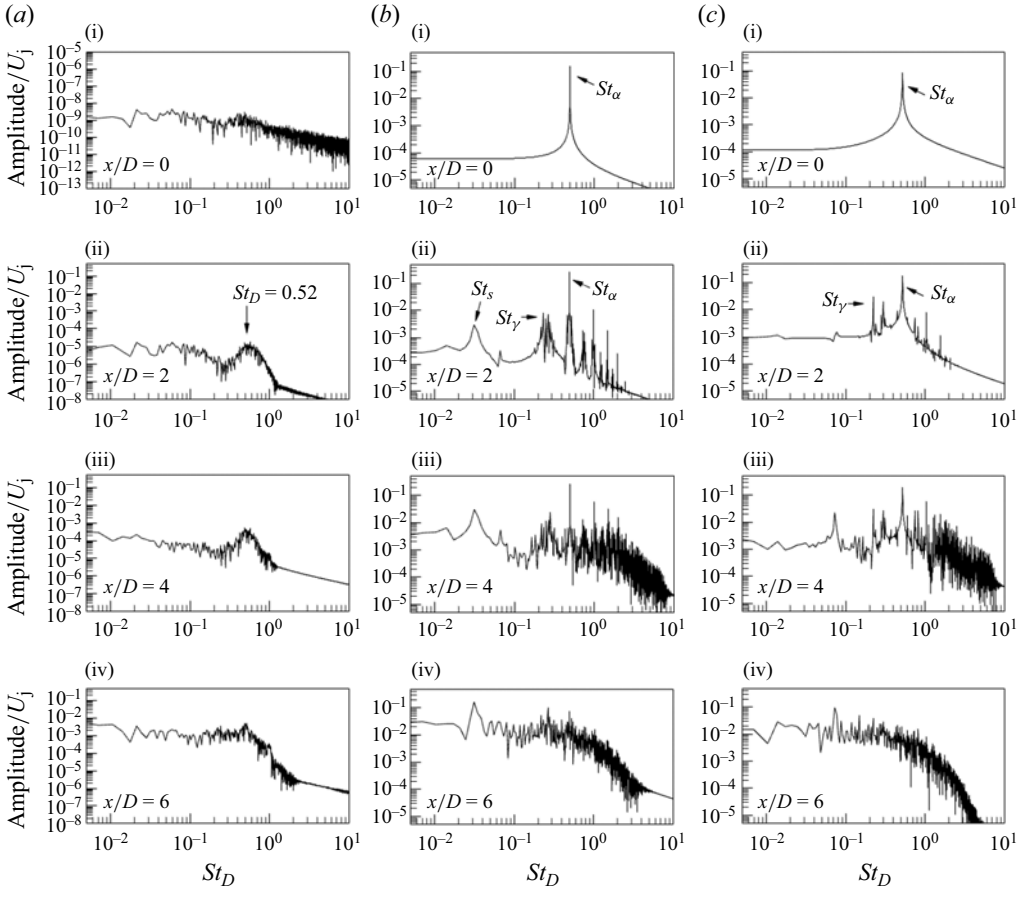


Figure 7. Axial velocity spectrum for the unforced (*a*-i–iii), the helical (*b*-i–iii) and the blooming jet (*c*-i–iii) at $x = 0D$, $2D$, $4D$ and $6D$.

BO by incorporating a deep-learning-enhanced surrogate model. This surrogate model combines the non-parametric method GP for fast local descent and the parametric method deep operator network (DeepONet) for efficient exploration of the search space. The proposed optimizer BO-DL is more efficient in searching for minima and more scalable to large datasets. To further understand the optimized high-dimensional solutions, we propose a topological analysis of the optimization data. The achieved control landscape features two persistent (pronounced) minima, a global minimum corresponding to a 7-armed blooming jet being generated, and a suboptimal parameter with a double-helix shape that performs comparably. Intriguingly, many of the less persistent minima also correspond to known actuated jet mixing mechanisms.

Compared with the unforced jet, both the helical and the blooming jet shorten the length of the potential core substantially from $7.5D$ to $2D$. The valley of the mean centreline velocity is located around $6D$ in the downwash, corresponding to the peak of the fluctuation profiles. The reversal flow in the blooming jet amplifies the local turbulence intensity, and leads to even negative velocity in the centreline. Both of the optimized control laws show the radial component dominates the non-axisymmetric forcing mode. The optimized forcing for a helical jet is a triple mode that combines the axisymmetric bulk component, a helical and a flapping mode in the periphery. The 7-armed blooming

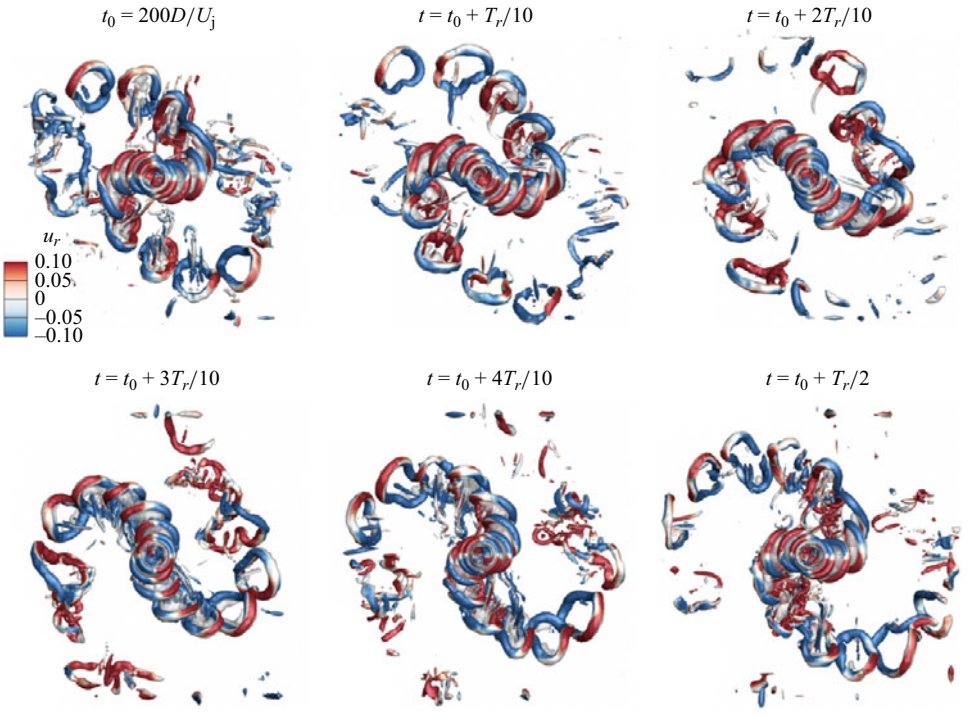


Figure 8. Instantaneous isosurfaces of Q-parameter ($Q = 0.5$) for the helical jet, coloured by the radial velocity, in a half-period of the arm rotation. Here, $T_r = 62.5D/U_j$.

jet is produced by a dual-mode forcing with only axisymmetric and helical modes. The better performance of the latter is attributed to the exact match between the axisymmetric forcing Strouhal number and the preferred mode frequency found by BO-DL. The forced flows are characterized by a distinct peak at the Strouhal number of the axisymmetric mode, and the effect of the helical forcing appears later. Intriguingly, a peak at the low Strouhal number in the helical jet coincides with the azimuthal motion of the helical arms.

This study emphasizes the importance of effective exploration for machine learning-based optimization in flow control, particularly in high-dimensional design spaces. The proposed BO-DL enhances the explorative feature of BO by improving the model accuracy and increasing the solvable model capacity. Therefore, BO-DL can serve as an alternative to classical BO when there is a need for greater complexity. In addition to parallelizing GP and DeepONet in the Bayesian framework, we can also incrementally increase the model complexity. For example, we can use the controller obtained by GP to accelerate the learning process of DeepONet. Furthermore, DeepONet can also be employed as a function approximator like deep reinforcement learning which deserves future study under a different framework – Bayesian experimental design. As an add on, the proposed persistent data topology analysis can help to characterize the control landscape from the discrete data produced by different optimizers. Persistent data minima indicate literature known and unknown mixing mechanisms. Finally, we expect the proposed BO-DL and topological data analysis for effective learning and characterization of the search spaces could contribute to more flow control problems.

Acknowledgements. We gratefully acknowledge inspiring and formative contributions of A. Blanchard and T. Sapsis to the employed BO algorithm and its interpretation. We express our sincere thanks to Z. Jiang for testing our optimization algorithms in a companion jet mixing experiment and deepening our physical insights

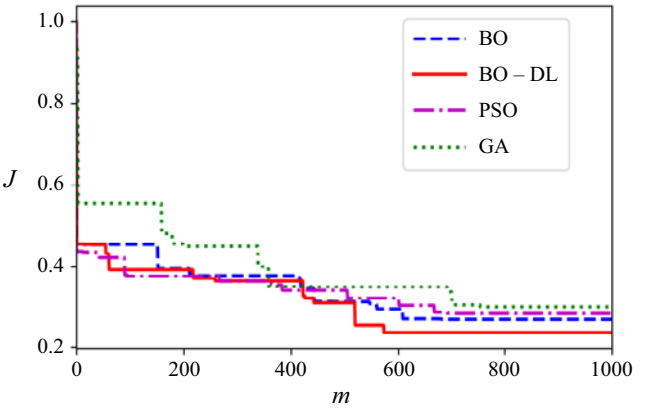


Figure 9. Learning curves of J_{min} using BO, BO-DL, PSO and GA.

and to T. Weise for enriching our roadmap on optimization algorithms and to H. Edelsbrunner and H.-C. Hege for a discussion of topological data analyses.

Funding. This work is supported by the National Natural Science Foundation of China under grants 12172109 and 12302293, by the Guangdong Basic and Applied Basic Research Foundation under grant 2022A1515011492 and by the Shenzhen Science and Technology Program under grant JCYJ20220531095605012. This work is also supported by the Polish National Science Center under grant 2018/31/B/ST8/00762.

Declaration of interests. The authors report no conflict of interest.

Author ORCIDs.
Yiqing Li https://orcid.org/0000-0003-2547-5363;
Bernd R. Noack https://orcid.org/0000-0001-5935-1962;
Tianyu Wang https://orcid.org/0000-0001-5331-1142;
Guy Y. Cornejo Maceda https://orcid.org/0000-0001-7499-7569;
Ethan Pickering https://orcid.org/0000-0002-4485-6359;
Tamir Shaqarin https://orcid.org/0000-0003-3327-3147;
Artur Tyliszczak https://orcid.org/0000-0002-4844-0289.

Appendix. Comparison of Bayesian optimization with two bio-inspired optimizers

The Bayesian optimizers, BO and BO-DL, employed in this study have been compared with two popular biologically inspired methods (Wahde 2008): PSO and GA. Here, a variant of PSO, particle swarm optimization through targeted, position-mutated, elitism is employed (Shaqarin & Noack 2023), and GA is realized following Wright (1991). Figure 9 displays the learning curve of one realization of the four methods. The learning curves give an indication of the learning speed of each method. Particle swarm optimization converges to a solution with the cost $J = 0.283$ slightly higher than BO ($J = 0.273$), and GA ends with a even higher cost $J = 0.3$. Figure 10 presents all the evaluated points (a), and the local minima (b) derived from the combined database. The tested solutions during the search are denoted by the coloured dots in figure 10(a). The derived minima are denoted by the filled circles with corresponding colours in figure 10(b). The converged solutions are depicted by stars; PSO and GA fall into the local minima in the upper left and the upper right corner, respectively. Interestingly, the search process of these methods show different features. Particle swarm optimization moves all the particles (magenta dots)

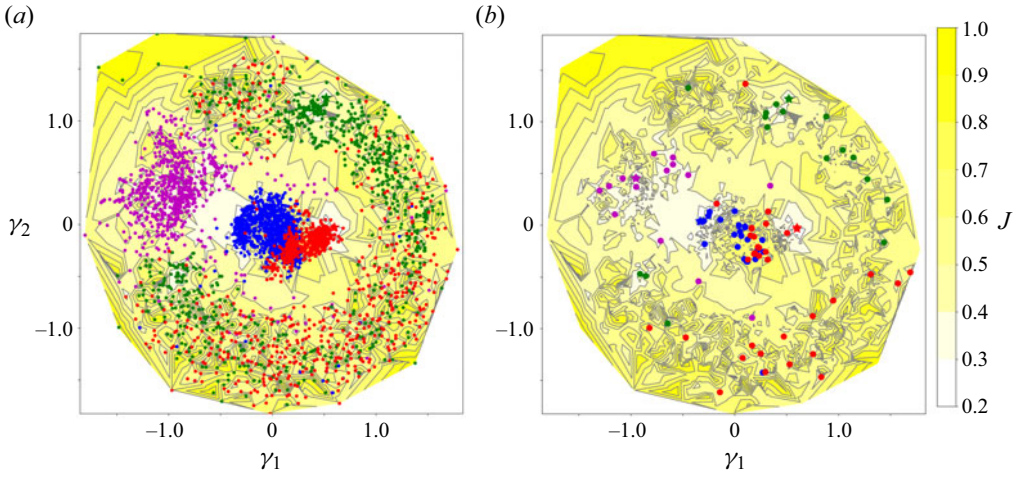


Figure 10. Landscape with visited data (*a*) and derived local minima (*b*) of BO (blue), BO-DL (red), PSO (magenta) and GA (green) from unknown in the whole space.

towards the best region detected. Finally, all particles accumulate in the upper left region and get stuck. The minima (magenta circles) are found along the direction of gradient descent. The genetic algorithm searches the minima in one neighbourhood but is extremely inefficient in exploring further regions. Most of the exploration away from the global minimum in the right upper region in figure 9(*a*) ends with no local minima in figure 9(*b*). Owing to the prediction by GP, BO converges to the region with a lower cost quickly. Moreover, with the deep-learning-enhanced surrogate model, BO-DL not only obtains the best minimum (red star) but also reveals more potential minima in a wider neighbourhood (red circles). In high-dimensional search spaces, exploration based on accurate estimators is more efficient than random exploration. For the current problem, the optimizer based on a surrogate model, particularly a deep-learning model, shows more advantages than bio-inspired optimizers.

References

Ball, C.G., Fellouah, H. & Pollard, A. 2012 The flow field in turbulent round free jets. *Prog. Aerosp. Sci.* **50**, 1–26.

Blanchard, A. & Sapsis, T. 2021 Bayesian optimization with output-weighted optimal sampling. *J. Comput. Phys.* **425**, 109901.

Blanchard, A.B., Cornejo Maceda, G.Y., Fan, D., Li, Y., Zhou, Y., Noack, B.R. & Sapsis, T.P. 2021 Bayesian optimization for active flow control. *Acta Mechanica Sin.* **37**, 1786–1798.

Boguslawski, A., Wawrzak, K. & Tyliszczak, A. 2019 A new insight into understanding the crow and champagne preferred mode: a numerical study. *J. Fluid Mech.* **869**, 385–416.

Brunton, S.L., Noack, B.R. & Koumoutsakos, P. 2020 Machine learning for fluid mechanics. *Annu. Rev. Fluid Mech.* **52**, 477–508.

Corke, T.C. & Kusek, S.M. 1993 Resonance in axisymmetric jets with controlled helical-mode input. *J. Fluid Mech.* **249**, 307–336.

Cornejo Maceda, G.Y., Li, Y., Lusseyran, F., Morzyński, M. & Noack, B.R. 2021 Stabilization of the fluidic pinball with gradient-based machine learning control. *J. Fluid Mech.* **917**, A42.

Crow, S.C. & Champagne, F.H. 1971 Orderly structure in jet turbulence. *J. Fluid Mech.* **48** (3), 547–591.

Danaila, I. & Boersma, B.J. 2000 Direct numerical simulation of bifurcating jets. *Phys. Fluids* **12** (5), 1255–1257.

Duriez, T., Brunton, S.L. & Noack, B.R. 2017 *Machine Learning Control – Taming Nonlinear Dynamics and Turbulence*. Fluid Mechanics and Its Applications, vol. 116. Springer.

EDELSBRUNNER, H. & HARER, J. 2008 Persistent homology – a survey. *Contemp. Maths* **453** (26), 257–282.
GOHIL, T.B., SAHA, A.K. & MURALIDHAR, K. 2015 Simulation of the blooming phenomenon in forced circular jets. *J. Fluid Mech.* **783**, 567–604.
GUASTONI, L., RABAULT, J., SCHLATTER, P., AZIZPOUR, H., & VINUESA, R. 2023 Deep reinforcement learning for turbulent drag reduction in channel flows. *Eur. Phys. J.* E **46**, 27.
GUTMARK, E. & HO, C.M. 1983 Preferred modes and the spreading rates of jets. *Phys. Fluids* **26**, 2932–2938.
HILGERS, A. & BOERSMA, B.J. 2001 Optimization of turbulent jet mixing. *Fluid Dyn. Res.* **29** (6), 345.
HUSSAIN, A.K.M.F. & ZAMAN, K.B.M.Q. 1980 Vortex pairing in a circular jet under controlled excitation. Part 2. Coherent structure dynamics. *J. Fluid Mech.* **101** (3), 493–544.
JORDAN, P. & COLONIUS, T. 2013 Wave packets and turbulent jet noise. *Annu. Rev. Fluid Mech.* **45** (1), 173–195.
KEMPF, A., KLEIN, M. & JANICKA, J. 2005 Efficient generation of initial- and inflow-conditions for transient turbulent flows in arbitrary geometries. *Flow Turbul. Combust.* **74**, 67–84.
KOUMOUTSAKOS, P., FREUND, J. & PAREKH, D. 2001 Evolution strategies for automatic optimization of jet mixing. *AIAA J.* **39** (5), 967–969.
LEE, M. & REYNOLDS, W.C. 1985 Bifurcating and blooming jets. *Tech. Rep.*. Thermosciences Division, Department of Mechanical Engineering, Stanford University.
LU, L., JIN, P., PANG, G., ZHANG, Z. & KARNIADAKIS, G.E. 2021 Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nat. Mach. Intell.* **3** (3), 218–229.
MANKBADI, R. & LIU, J.T.C. 1981 A study of the interactions between large-scale coherent structures and fine-grained turbulence in a round jet. *Phil. Trans. R. Soc.* A **298** (1443), 541–602.
NAIR, N.J. & GOZA, A. 2023 Bio-inspired variable-stiffness flaps for hybrid flow control, tuned via reinforcement learning. *J. Fluid Mech.* **956**, R4.
NATHAN, G.J., MI, J., ALWAHABI, Z.T., NEWBOLD, G.J.R. & NOBES, D.S. 2006 Impacts of a jet's exit flow pattern on mixing and combustion performance. *Prog. Energy Combust. Sci.* **32** (5), 496–538.
PAREKH, D.E. 1989 *Bifurcating Jets at High Reynolds Numbers*. Stanford University.
PICKERING, E., GUTH, S., KARNIADAKIS, G.E. & SAPSIS, T.P. 2022 Discovering and forecasting extreme events via active learning in neural operators. *Nat. Comput. Sci.* **2** (12), 823–833.
PINO, F., SCHENA, L., RABAULT, J. & MENDEZ, M.A. 2023 Comparative analysis of machine learning methods for active flow control. *J. Fluid Mech.* **958**, A39.
RABAULT, J., KUCHTA, M., JENSEN, A., RÉGLADE, U. & CERARDI, N. 2019 Artificial neural networks trained through deep reinforcement learning discover control strategies for active flow control. *J. Fluid Mech.* **865**, 281–302.
SADEGHI, H. & POLLARD, A. 2012 Effects of passive control rings positioned in the shear layer and potential core of a turbulent round jet. *Phys. Fluids* **24** (11), 115103.
SHAABANI-ARDALI, L., SIPP, D. & LESSHAFFT, L. 2020 Optimal triggering of jet bifurcation: an example of optimal forcing applied to a time-periodic base flow. *J. Fluid Mech.* **885**, A34.
SHAHRIARI, B., SWERSKY, K., WANG, Z., ADAMS, R.P. & DE FREITAS, N. 2015 Taking the human out of the loop: a review of Bayesian optimization. *Proc. IEEE* **104** (1), 148–175.
SHAQARIN, T. & NOACK, B.R. 2023 A fast-converging particle swarm optimization through targeted, position-mutated, elitism (PSO-TPME). *Intl J. Comput. Intell. Syst.* **16**, 6.
DA SILVA, C.B. & MÉTAIS, O. 2002 Vortex control of bifurcating jets: a numerical study. *Phys. Fluids* **14** (11), 3798–3819.
SONODA, T., LIU, Z., ITOH, T. & HASEGAWA, Y. 2023 Reinforcement learning of control strategies for reducing skin friction drag in a fully developed turbulent channel flow. *J. Fluid Mech.* **960**, A30.
SUZUKI, H., KASAGI, N. & SUZUKI, Y. 1999 Active control of an axisymmetric jet with an intelligent nozzle. In *First Symposium on Turbulence and Shear Flow Phenomena*. Begell House.
TYLISZCZAK, A. 2014 A high-order compact difference algorithm for half-staggered grids for laminar and turbulent incompressible flows. *J. Comput. Phys.* **276**, 438–467.
TYLISZCZAK, A. 2015*a* LES-CMC study of an excited hydrogen flame. *Combust. Flame* **162** (10), 3864–3883.
TYLISZCZAK, A. 2015*b* Multi-armed jets: a subset of the blooming jets. *Phys. Fluids* **27** (4), 041703.
TYLISZCZAK, A. 2018 Parametric study of multi-armed jets. *Intl J. Heat Fluid Flow* **73**, 82–100.
TYLISZCZAK, A. & GEURTS, B.J. 2014 Parametric analysis of excited round jets-numerical study. *Flow Turbul. Combust.* **93**, 221–247.
UTKIN, Y.G., KESHAV, S., KIM, J.H., KASTNER, J., ADAMOVICH, I.V. & SAMIMY, M. 2006 Development and use of localized arc filament plasma actuators for high-speed flow control. *J. Phys.* D*: Appl. Phys.* **40** (3), 685.

VIGNON, C., RABAULT, J., VASANTH, J., ALCÁNTARA-ÁVILA, F., MORTENSEN, M. & VINUESA, R. 2023*a* Effective control of two-dimensional Rayleigh–Bénard convection: invariant multi-agent reinforcement learning is all you need. *Phys. Fluids* **35** (6), 065146.

VIGNON, C., RABAULT, J. & VINUESA, R. 2023*b* Recent advances in applying deep reinforcement learning for flow control: perspectives and future directions. *Phys. Fluids* **35** (3), 031301.

WAHDE, M. 2008 *Biologically Inspired Optimization Methods: An Introduction*. WIT Press.

WANG, T., CORNEJO MACEDA, G.Y. & NOACK, B.R. 2023*a* *XPDT: A Toolkit for Persistent Data Topology*. Universitätsbibliothek der Technischen Universität Braunschweig.

WANG, T., YANG, Y., CHEN, X., LI, P., IOLLO, A., CORNEJO MACEDA, G.Y. & NOACK, B.R. 2023*b* Topologically assisted optimization for rotor design. *Phys. Fluids* **35** (5), 055105.

WAWRZAK, K., BOGUSLAWSKI, A. & TYLISZCZAK, A. 2015 LES predictions of self-sustained oscillations in homogeneous density round free jet. *Flow Turbul. Combust.* **95**, 437–459.

WILLIAMS, C.K. & RASMUSSEN, C.E. 2006 *Gaussian Processes for Machine Learning*. MIT Press.

WRIGHT, A.H. 1991 Genetic algorithms for real parameter optimization. In *Foundations of Genetic Algorithms*, vol. 1, pp. 205–218. Elsevier.

XU, D. & ZHANG, M. 2023 Reinforcement-learning-based control of convectively unstable flows. *J. Fluid Mech.* **954**, A37.

ZHOU, Y., FAN, D., ZHANG, B., LI, R. & NOACK, B.R. 2020 Artificial intelligence control of a turbulent jet. *J. Fluid Mech.* **897**, A27.