SUPPORTING DOCUMENTATION

OSCILLATORY FLOW ASSESSED

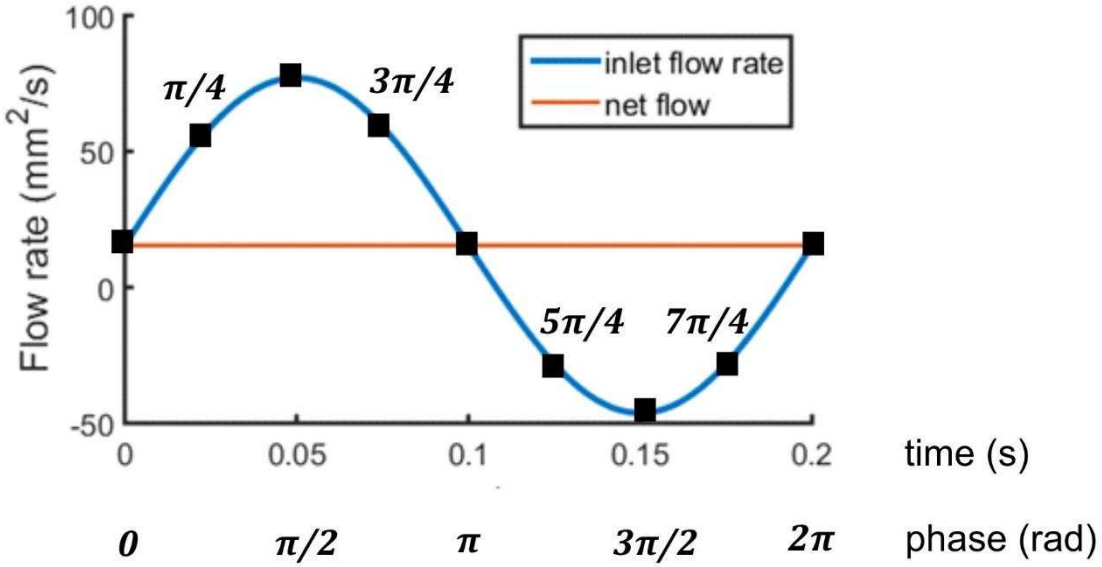

Figure S1 Inlet flow rate for the 2D mCOBR CFD simulation. The net flow is also plotted for comparison. Black squares represent the points in the cycle for which solutions were computed.

CFD VALIDATION

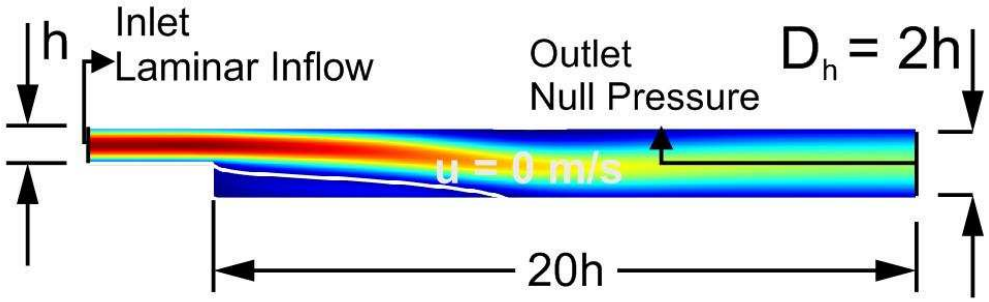

Figure S2 CFD Model for a backward facing step problem. The step size, h, is selected to be 5mm, following Armaly's experiment. The computational domain is chosen to be long enough to neglect exit effects. Shown in white, the line for which the axial velocity, u is null. The distance between the step and the intersection of this line with the lower wall is the reattachment length. Colormap shows the velocity magnitude for the case Re=400. The mean inlet velocity is computed from the Reynolds number selected following

$$Re = \frac{\rho u_{avg} D_h}{\mu}$$

for every Re case assessed.

As figures S1 and S2 show, the velocity profiles also exhibit good agreement between the experimental and computational data.

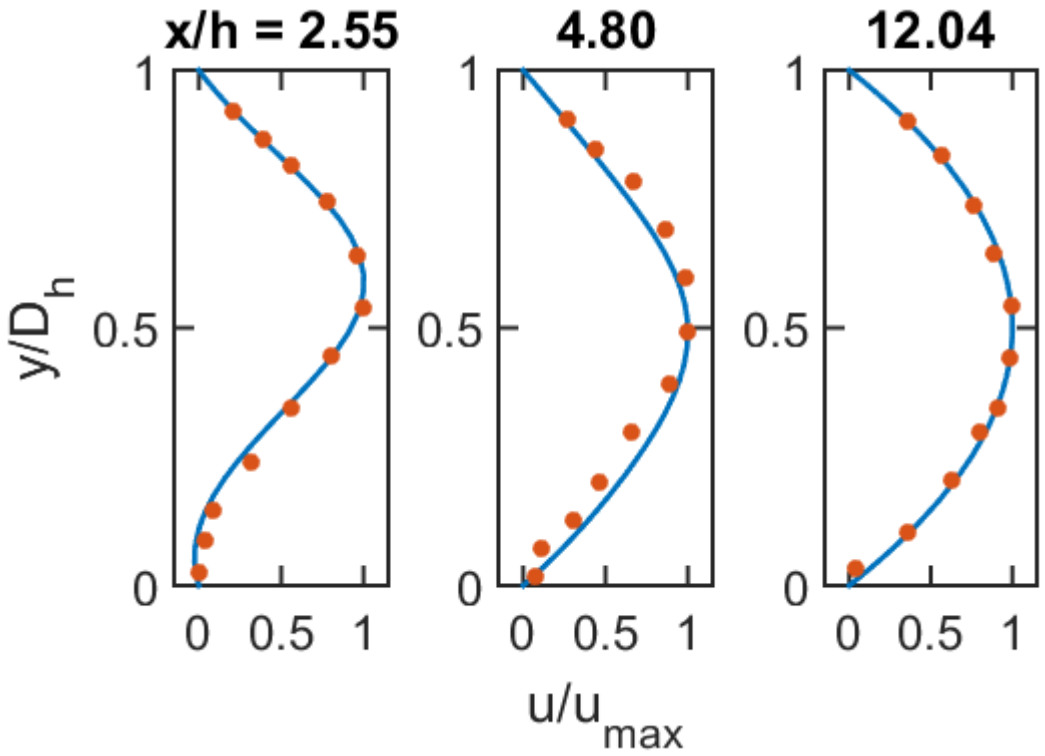

Figure S3 Axial velocity profiles for three different sections downstream of the step for Re = 100. Height is normalised by the height of the channel and velocities by the maximum velocity for better comparison. Blue line: CFD results. Red dots: Armaly's experimental data.

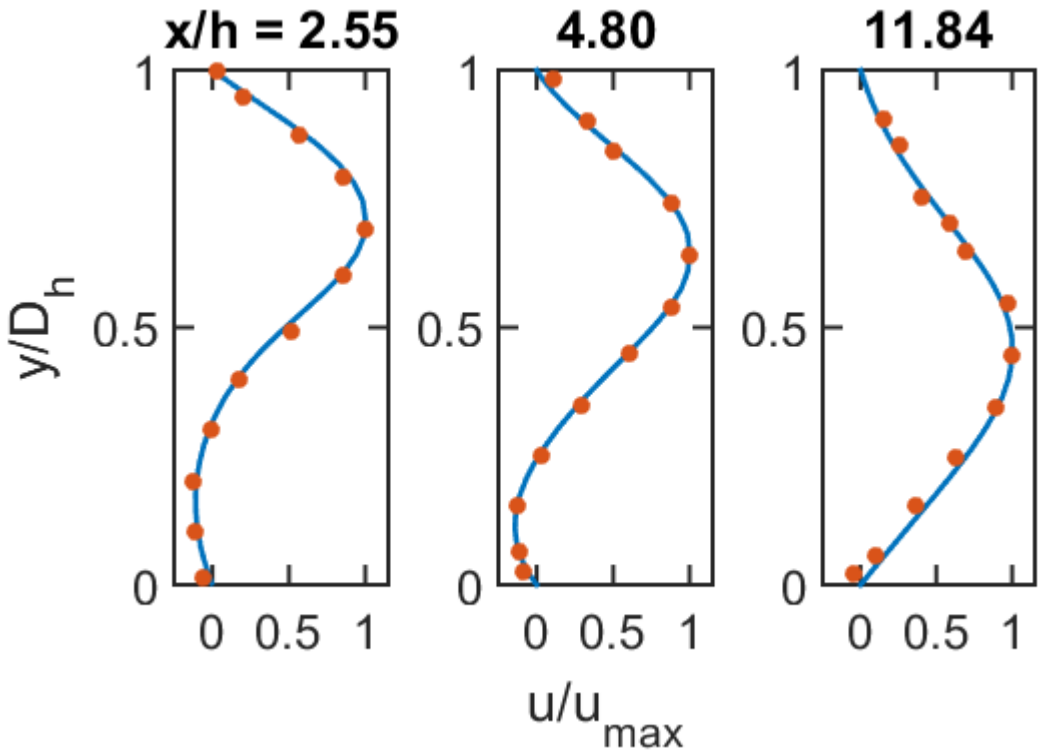

Figure S4 Axial velocity profiles for three different sections downstream of the step for Re = 389. Height is normalised by the height of the channel and velocities by the maximum velocity for better comparison. Red dots: Armaly's experimental data.

In order to compare the meshes employed by the solver, a quarter of a COBR cell is compared with the equivalent geometry in the backward facing step. These geometries are different in size, but the relationships of the region of interest features with the step size (baffle height in the case of the OBR) are kept the same as shown in Figure S4.

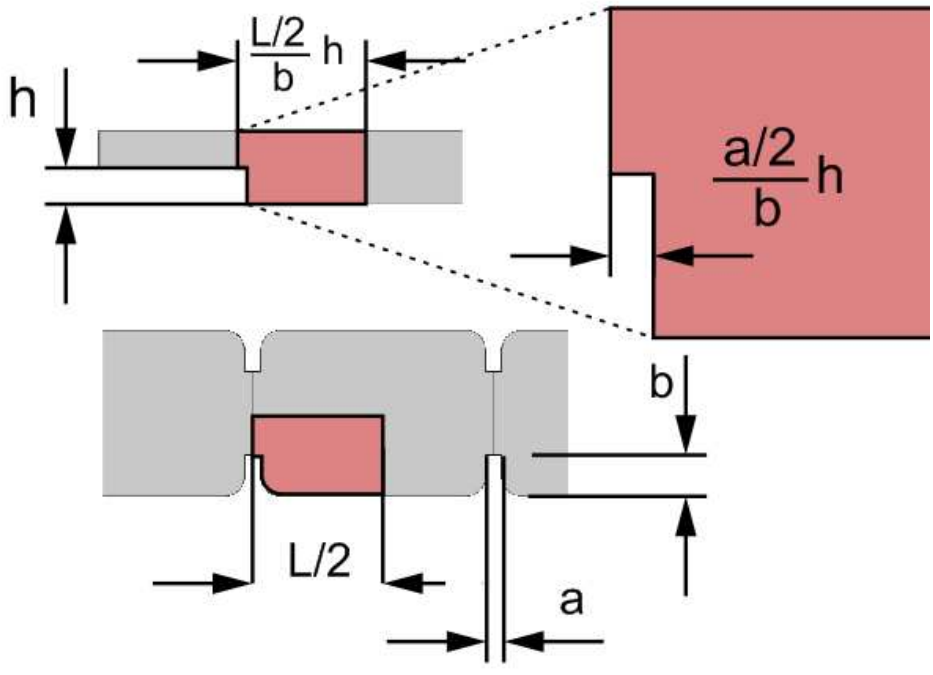

Figure S5 Region of interest selected for comparison of mesh densities for the backward facing step and COBR geometries.

The number of mesh elements for the highlighted area in the backward facing step experiment is 1312 compared with 1237 in the case of the quarter of COBR cell. The quadrangular elements of the top boundary layer for the backward facing step were discounted, as this boundary does not exist in the case of the COBR. The similarity of mesh densities is high, with less than 6% difference in the number of elements generated by COMSOL's predefined physics-controlled 'fine' mesh.

Comparing the Reynolds number in the OBR, defined as

$$Re = \frac{\rho u_{avg} {}^{D}/_{2}}{\mu}$$

where D is the cell width and D/2 is proportional to D_h. Taking this into account, the highest Reynolds number reached for a mCOBR simulation across all cases evaluated is 129. Taking a look at figure 7, this is well below 500 and in the regime considered to fit to a high degree with the CFD simulations.

SURROGATE MODEL CROSS-VALIDATION

Table S1 Cross-validation results. Sample size: 27 points.

	K left out	**θ**	**CV RMSE**	**Training RMSE**	**CV σ^2**	**Test RMSE**
η	1 (LOO)	10	0.0533	0.0013	$29.3405*10^{-4}$	0.0683
	3	10	0.0535	0.0013	$11.2370*10^{-4}$	0.0683
	5	2.07	0.0559	0.0207	$7.4275*10^{-4}$	0.0627
	7	1.34	0.0589	0.0266	$6.9889*10^{-4}$	0.0606
	9	0.90	0.0646	0.0310	$7.5544*10^{-4}$	0.0593
		0		0.0419		0.05651
σ^2	LOO	3.40	0.0277	0.0073	$7.9538*10^{-4}$	0.0171
	3	2.76	0.0273	0.0090	$2.9719*10^{-4}$	0.0175
	5	2.58	0.0283	0.0096	$2.0285*10^{-4}$	0.0176
	7	2.13	0.0318	0.0111	$2.1892*10^{-4}$	0.0180
	9	1.69	0.0367	0.0129	$2.4812*10^{-4}$	0.0185
		10		0.0006		0.0164

Table S1 contains the following information. CV RMSE is the average RMSE over all $k_{left\ out}$ points that are removed from the full training set of n=27 DoE points,

obtained using a surrogate model trained on the remaining (n- $k_{left\ out}$) DoE points. This value is minimised by the golden search algorithm that finds θ. Training RMSE represents the RMSE calculated over all 27 DoE points using the surrogate model already tuned and Test RMSE refers to the RMSE calculated over the nested DoE points held out during the LOO and MC cross-validation procedures. This means that the learning algorithm does not have access to the hold out dataset and hence, test RMSE is indicative of how the tuned model will perform for uncalibrated data. In each case θ represents the optimal value that leads to the smallest associated CV RMSE.

For the η surrogate model, Table S1 shows how, as $k_{left\ out}$ increases, both CV RMSE and Training RMSE also increase. This loss in precision is a result of the surrogate model learning with a lower number of points. Conversely, Test RMSE decreases as the number of points left out grows. This is an indicator of the model over-fitting the training set when only a small number of points are left out [1, 2]. Even when $k_{left\ out}$ = 9, a substantial difference between test RMSE and training RMSE values indicates the model is still over-fitting. In order to prevent this behaviour, and following the trend in decreasing values of θ, a surrogate model for θ=0, which corresponds to the conventional least squares method with 10 constant coefficients in the quadratic regression for the 3 design variables was computed. This yielded the lowest value of the test RMSE and also the minimum difference between test and training RMSE.

Similar behaviour is seen for the σ^2 surrogate model, for which the cross-validation methods find decreasing values of theta as $k_{left\ out}$ grows. Apart from the cross-validated surrogates, an extra one is computed for θ=10, yielding the lowest values for both training and test RMSE. However, as discussed before, the huge gap between these values suggests a severe over-fitting of the training data. Following [3] the most appropriate $k_{left\ out}$ is chosen which minimises the ratio of the training RMSE to the test RMSE, so the surrogate models used in the following optimisation study use MCCV with $k_{left\ out}$=9 and θ=1.6898.

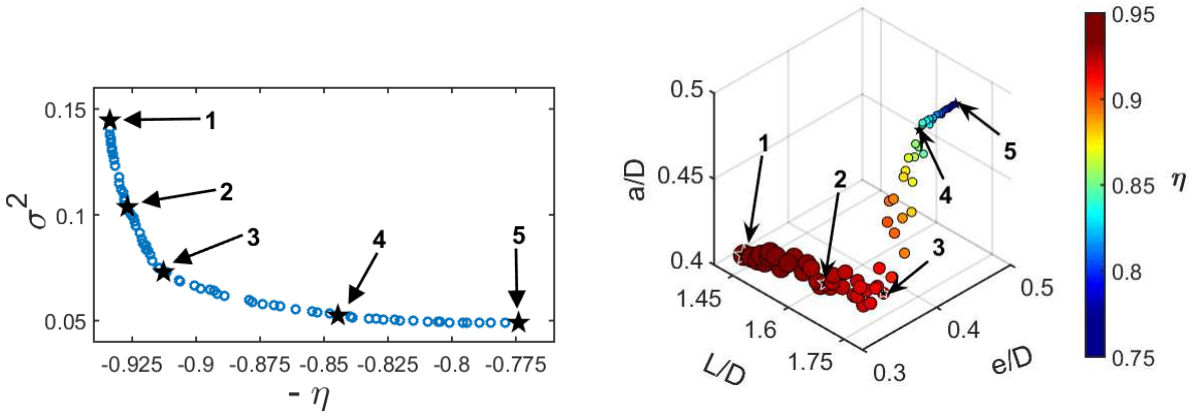

Figure S6. Pareto front (left) and its points represented in the design space (right) where the colour indicates the value of η and the area of the circles is proportional to σ^2. The validation points evaluated are indicated as black stars.

[1] M. Kuhn and K. Johnson, Applied predictive modeling. New York: Springer, 2013.

[2] G. James, D. Witten, T. Hastie, and R. Tibshirani, An introduction to statistical learning : with applications in R, 2013.

[3] "Bias Variance Decomposition," in Encyclopedia of Machine Learning, C. Sammut and G. I. Webb, Eds., ed Boston, MA: Springer US, 2010, pp. 100-101.