AIAA SciTech Forum
9 - 13 January 2017, Grapevine, Texas
19th AIAA Non-Deterministic Approaches Conference

10.2514/6.2017-1090

Check for updates

Developing Design Insight Through Active Subspaces

Zachary R. del Rosario* and Gianluca Iaccarino†
Stanford University, Stanford, CA, 94305, USA
Paul Constantine‡
Colorado School of Mines, Golden, CO 80401, USA

Engineering design is the iterative process of defining an engineering specification; combining form and function to achieve a performance goal. It is a process which demands, among other things, *design insight* – a conceptual understanding of the physical processes to be manipulated and leveraged. Such insights guide the conceptual phases of design, encourage novel solutions, reduce the search space (i.e. the dimensionality of the problem), and lead to confidence in the final result. In this work, we focus on developing design insight through modern dimension reduction tools; namely, Active Subspaces.

Nomenclature

Active Subspace Formulation

$\boldsymbol{C}$	Average outer product of gradients
$\boldsymbol{\Lambda}$	Eigenvalues of $\boldsymbol{C}$
$\boldsymbol{\Lambda}_1$	Active eigenvalues
$\boldsymbol{\Lambda}_2$	Inactive eigenvalues
$\boldsymbol{W}$	Eigenvectors of $\boldsymbol{C}$
$\boldsymbol{W}_1$	Active directions
$\boldsymbol{W}_2$	Inactive directions
$\rho(\boldsymbol{x})$	Weight function
$\boldsymbol{y}$	Active variables
$\boldsymbol{z}$	Inactive variables
k	Dimensionality of Active Subspace

Design Problem

$\mathcal{X}$	Input parameter space
$\boldsymbol{D}$	Dimension matrix
∇f	Gradient
π	Dimensionless parameter
ψ	Physical law via Buckingham π
$\boldsymbol{x}$	Input parameters
$f(\boldsymbol{x})$	Quantity of interest
h	Physical law as ridge function
l	Number of independent dimensions
m	Dimensionality of input space
n	Number of dimensionless parameters

I. Introduction

THE objectve of engineering design is to generate feasible specifications for an engineering solution. Beitz and Pahl[1] describe it thusly; "Design is an engineering activity that... provides the prerequisites for the physical realization of solution ideas." This is a process which balances two competing effects: To achieve 'physical realization' the design must comply with physical laws, while to be attractive in terms of performance or cost it must also be innovative. If a design is already fully and widely understood, then it is highly unlikely to be novel! The enterprise of engineering design leverages the knowledge of past projects, but also requires new *insight* – deep understanding of the processes at work, which guides the process of developing novel solutions.

As an illustrative example, consider the problem of designing a wing, with a focus on lift. An aerodynamicist will likely recall the simplified case of the thin airfoil,[2] and will quickly focus on changes of the camber line and angle of attack. The aerodynamicist has design insight into the problem, and has performed a kind of *dimension reduction* on the problem: Rather than considering the infinite-dimensional space of all possible geometrical perturbations of the wing geometry and operating scenarios, she knows that for the quantity of interest (QoI) in question (Lift) and physical problem (inviscid, incompressible thin airfoil), there are only

*PhD Candidate, Aeronautics and Astronautics, Address, AIAA Student Member.
†Associate Professor, Department of Mechanical Engineering, AIAA Associate Fellow
‡Ben L. Fryrear Assistant Professor, Applied Mathematics and Statistics, AIAA Member

Copyright © 2017 by the American Institute of Aeronautics and Astronautics, Inc. All rights reserved.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

two relevant variables (inputs) to consider (angle of attack and camber). The focus on just two inputs can seem constrictive when moving from simple cases to questions of more practical interest; for example, the design of multi-component wing systems for a commercial airplane. However, the insights developed from thin airfoil theory remain largely relevant: One may regard flaps and slats as devices that modify the effective camber of an airfoil,[2] and variable-camber systems are a potential means to achieve high-lift configurations.[3] Of course, one must check these insights when moving to seemingly related physical cases. The angle of attack remains important for *supersonic* thin airfoils, but the dependence on the camber line disappears.[4] This example illustrates a few important points:

- Design insights guide the development of novel solutions.
- Such insights also lead to dimension reduction.
- These insights must be tested when moving to related problems.

Classical dimensional analysis (DA) provides a framework to reduce the complexity of an engineering problem and to gain insight into the governing physics. The main tool – the Buckingham π theorem – provides explicit guidelines on the number of non-dimensional groups controlling independent variables relevant in the physical problem. However, no quantitative information can be extracted to rank the importance of various inputs depending on specific operating scenarios. Moreover, dimensional analysis cannot distinguish between inputs having the same units without further physical knowledge. In the airfoil case discussed above, DA might identify a lengthscale as an important design characteristic, but cannot recognize the thickness or the chord of the airfoil as being the critical one. Finally, Buckingham π alone cannot determine when important parameters as missing from the analysis, i.e. when there are *hidden parameters*.

In this work, we consider the problem of generating and testing design insights. Since these insights are closely tied to reducing the dimensionality of a system, we approach this problem from the perspective of dimension reduction, and leverage the Active Subspace technique.[5] Previous studies with this technique have focused on optimization through surrogate models,[6] uncertainty quantification,[7] or sensitivity analysis and dimension reduction.[8] Instead, we focus on interpreting the Active Subspace, recovering classical insights in well-known engineering problems, and demonstrating new insights. One important contribution of the paper is also the clarification of the link between Active Subspaces and Dimensional Analysis.

Section II introduces the Active Subspace method. Section III combines this technique with Classical Dimensional Analysis, and draws a connection between Active Subspaces and the Buckingham π Theorem. We demonstrate this connection on the physical problem of pipe flow, and identify and rank dimensionless parameters which govern the system – *active dimensionless parameters*. Section IV addresses an important failure mode of Classical Dimensional Analysis; that of missing an important governing parameter in our analysis. We present a method for detecting such *hidden parameters*, and demonstrate this technique on our pipe flow example. Section V considers *shape design*, and the identification of important deformations via the Active Subspace. We demonstrate this approach in the context of airfoil design.

II. Active Subspaces

This section summarizes and motivates the Active Subspace method; further details may be found in Constantine.[5] Suppose we have some QoI $f(\boldsymbol{x})$ over some domain $\mathcal{X} \subseteq \mathbb{R}^m$. As m increases, the computational cost of studying our QoI grows alarmingly quickly (Fig. 1). This *curse of dimensionality* is the motivating issue for Active Subspaces. Informally, the Active Subspace gives us a linear subspace decomposition of $\mathbb{R}^m$, split into *active* and *inactive* directions. The idea is to either ignore or treat with lower fidelity the inactive directions, and retain the active ones; this effectively reduces m. Note that while this allows us to reduce dimensionality, it also provides some design insight into the problem. The Active Subspace gives us those directions in parameter space along which the QoI changes most, on average. Thus, these directions are of great interest to a designer.

Formally, we scale and shift our parameters $\boldsymbol{x}$ such that the domain is the hypercube $\mathcal{X} = [-1, 1]^m$, assume the gradient ∇f exists, and define a weight function $\rho(\boldsymbol{x})$. This ρ may represent a probabilistic description of the input variables (i.e. a joint density function), or simply a function which weights regions of parameter space according to the designer's interest[a]. We then compute the following matrix

[a]Subject to the constraint $\rho(\boldsymbol{x}) > 0$ for $x \in \mathcal{X}$, and $\rho = 0$ for $x \notin \mathcal{X}$.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

$$\boldsymbol{C} = \int_{\mathbb{R}^m} (\nabla f(\boldsymbol{x}))(\nabla f(\boldsymbol{x}))^T \rho(\boldsymbol{x})\, d\boldsymbol{x}, \qquad (1)$$

where $d\boldsymbol{x} = \prod_{i=1}^{m} dx_i$; thus Equation 1 is an m-dimensional integral. Since $\boldsymbol{C}$ is symmetric positive-definite, it admits an eigenvalue decomposition $\boldsymbol{C} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^T$. We then choose a split based on the eigenvalues of $\boldsymbol{C}$, that is

$$\boldsymbol{W} = [\boldsymbol{W}_1\, \boldsymbol{W}_2],\ \boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_1 & \\ & \boldsymbol{\Lambda}_2 \end{bmatrix} \qquad (2)$$

where $\boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_1, \ldots, \lambda_k)$ with $k < m$, and $\boldsymbol{W}_1$ contains the first k eigenvectors. If $\boldsymbol{C}$ is estimated, say via some sampling method, then it is recommended for the purposes of accuracy that the split be chosen to exploit eigenvalue gaps.[5] With a split of $\boldsymbol{C}$ chosen, we define the active ($\boldsymbol{y}$) and inactive ($\boldsymbol{z}$) variables

$$\boldsymbol{y} = \boldsymbol{W}_1^T \boldsymbol{x} \in \mathbb{R}^k,\ \boldsymbol{z} = \boldsymbol{W}_2^T \boldsymbol{x} \in \mathbb{R}^{m-k}. \qquad (3)$$

In this way, the Active Directions $\boldsymbol{W}_1$ define linear combinations of the input parameters. One can then treat the Inactive Variables $\boldsymbol{z}$ with lower fidelity, focusing computational effort on the active directions. As an illustrative example, consider the function $f(\boldsymbol{x}) = \frac{1}{2}(.7x_1 + .3x_2)^2$. Clearly, the quantity of interest f changes along the direction $[.7, .3]^T$, and does not change at all along $[.3, -.7]^T$ (Fig. 2). The Active Subspace method will identify these directions, both allowing us to reduce the dimensionality of f, and providing some insight into how changes in parameter space manifest in f.

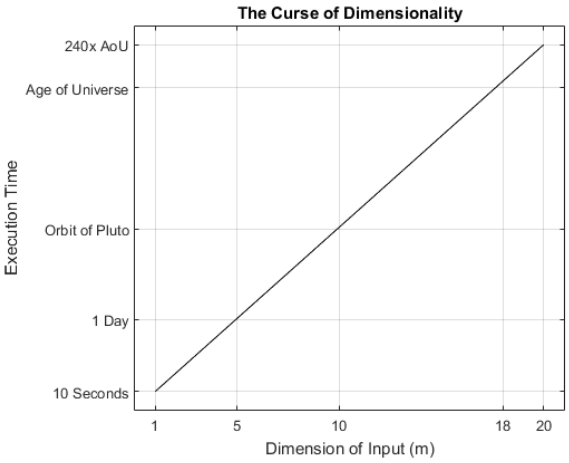

Figure 1. The computational cost of parameter studies increases exponentially with the dimensionality of the inputs; here we use a simple heuristic of sampling 10 points per dimension. If a computer code implementing f executes in one second, then the total execution time is 10^m seconds. This quickly becomes infeasible for modestly large m.

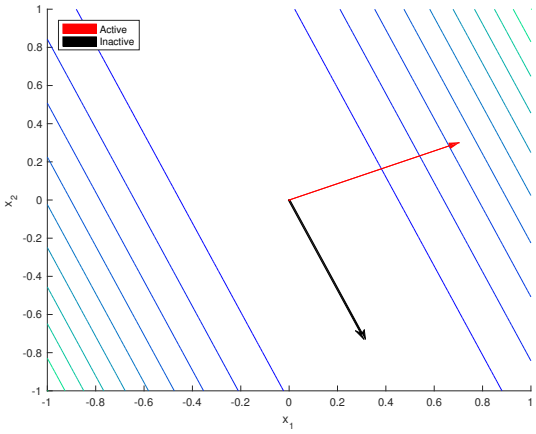

Figure 2. Active and Inactive directions for example function. Note that on average, the function changes more along the active directions than the inactive directions; in this case, the change in the inactive direction is exactly zero.

III. Dimensional Analysis and Active Subspaces

In this section we draw connections between Dimensional Analysis and Active Subspaces. We begin with a description of our example physical problem – that of pipe flow. Then we present a review of Classical Dimensional Analysis, and apply the technique to our example problem. We then draw connections between Dimensional Analysis and the Active Subspace, augmenting the classical technique with additional computational power. The result is the identification and ranking of dimensionless parameters which govern our QoI – *active dimensionless parameters.*

A. Mathematical Model of Pipe Flow

Our first example system is the classical problem of viscous flow through a pipe. Here, we consider the problem of setting the bulk velocity V by choosing the physical parameters fluid density ρ and viscosity μ,

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

pipe diameter D and characteristic wall roughness ϵ, and the pressure gradient $\frac{\Delta P}{L}$. We non-dimensionalize our QoI to form the Reynolds number $\rho V D/\mu$; we may return to V easily by re-dimensionalizing. Using the Active Subspace, we will determine relevant dimensionless parameters, for different parameter regimes. These specific parameter bounds are given in the Appendix.

The diagram in Figure 3 plots the Reynolds number Re against the dimensionless pressure $\frac{\rho D^3}{\mu}\frac{\Delta P}{L}$ and relative roughness $\frac{\epsilon}{D}$. Below a critical Reynolds number, taken to be $Re_c = 3 \times 10^3$, the friction factor satisfies the Poiseuille relation[9]

$$f = \frac{64}{Re}, \tag{4}$$

while for $Re > Re_c$, the Colebrook equation[10] models the relationship between the friction factor and other quantities

$$\frac{1}{\sqrt{f}} = -2.0 \log_{10}\left(\frac{1}{3.7}\frac{\epsilon}{D} + \frac{2.51}{Re\sqrt{f}}\right). \tag{5}$$

This relationship is intended to model behavior through transition to full turbulence. Using Equations 4 and 5, we consider how Re depends on $\{\rho, \mu, D, \epsilon, \frac{\Delta P}{L}\}$. This necessitates solving for V in the expressions above, and normalizing by the appropriate input parameters to form Re. In practice, we choose between Equations 4 and 5 by computing Re_t based on the Colebrook equation, and choosing the turbulent branch if $Re_t > Re_c$, and the Poiseuille relation otherwise. We treat this model as a black box, interpreted as either the output of some numerical process for which we do not have a closed form solution, or as the result of a laboratory experiment.

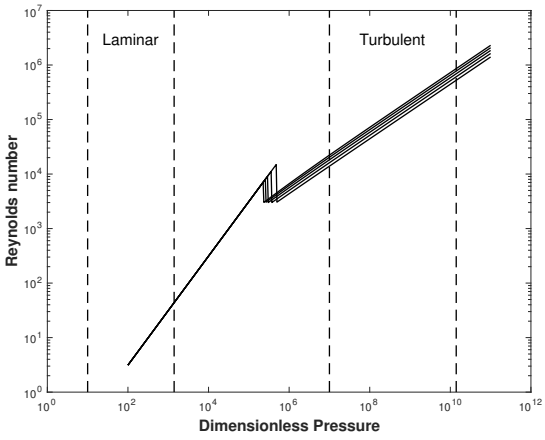

Figure 3. This diagram plots the Reynolds number against the dimensionless pressure. Different branches correspond to different values of Relative Roughness. Transition from laminar flow (Poiseuille relation) to turbulent flow (Colebrook equation) is assumed to begin at a critical Reynolds number of $Re_c \approx 3 \times 10^3$.

B. Classical Dimensional Analysis

Many physical systems of practical interest are controlled by variables with incommensurable units. This places certain constraints on the form that the governing physical laws can take, a fact made explicit in the Buckingham π theorem.[11] Assuming that we have a system with l independent dimensions and m input parameters $\{q_i\}_{i=1}^m$ with $m > l$, the Buckingham π theorem states that the physical law for a dimensionless QoI Π is governed by $n = m - l$ *dimensionless parameters*; symbolically

$$\pi = \psi(\pi_1, \ldots, \pi_n). \tag{6}$$

where each π_i is constructed from products of the dimensional parameters

$$\pi_j = \prod_{i=1}^{m} q_i^{\boldsymbol{V}_{ij}}, \tag{7}$$

such that their units cancel, i.e. they are dimensionless. The matrix $\boldsymbol{V} \in \mathbb{R}^{m \times n}$ is a set of powers chosen such that the π_j are dimensionless. The matrix $\boldsymbol{V}$ may be found by finding a basis for the nullspace of the *Dimension Matrix* $\boldsymbol{D}$; that is, the matrix which contains the fundamental units for each input parameter. Table 1 contains an example Dimension Matrix corresponding to our pipe flow example.

Computing the nullspace of $\boldsymbol{D}$ in the pipe flow example results in two dimensionless parameters ($n = 5 - 3 = 2$), an example of which would be ϵ/D and $\frac{\rho D^3}{\mu}\frac{\Delta P}{L}$, which we call the relative roughness and normalized pressure, respectively. Note that this choice is not unique; we could easily choose $\frac{\epsilon}{D}\frac{\rho D^3}{\mu}\frac{\Delta P}{L}$ and $\frac{\rho D^3}{\mu}\frac{\Delta P}{L}$ instead.

In the classical setting, this is where Dimensional Analysis ends: We have successfully reduced the dimensionality of the system from five dimensional inputs to just two dimensionless ones, but cannot say

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

	ρ	μ	D	ϵ	$\frac{\Delta P}{L}$
M	1	1	0	0	1
L	-3	-1	1	1	-2
T	0	-1	0	0	-2

Table 1. Dimension Matrix $\boldsymbol{D}$ for pipe flow example.

more without studying a model or data. In the case where we have access to such information, we can augment Dimensional Analysis with the Active Subspace procedure.

C. Dimensional Analysis with Active Subspaces

Note that the Active Subspace procedure defines Active Variables $\boldsymbol{y}$ which are *linear combinations* of our input parameters $\boldsymbol{q}$. From Equations 6 and 7, we can see that seeking linear combinations of the inputs in our pipe flow example is inadvisable – the Buckingham π theorem implies that we would be better off searching for *products* of the inputs. This can be accomplished by taking advantage of the properties of the log operator

$$\pi_j = \exp(\boldsymbol{v}_j^T \log(\boldsymbol{q})), \tag{8}$$

where the log and exp of a vector are understood to be elementwise operations. If we define new variables $\boldsymbol{x} = \log(\boldsymbol{q})$ and substitute Equation 8 into Equation 6, we find

$$\begin{aligned} \pi &= \psi(\exp(\boldsymbol{v}_1^T \boldsymbol{x}), \dots, \exp(\boldsymbol{v}_n^T \boldsymbol{x})), \\ \pi &= h(\boldsymbol{v}_1^T \boldsymbol{x}, \dots, \boldsymbol{v}_n^T \boldsymbol{x}), \\ \pi &= h(\boldsymbol{V}^T \boldsymbol{x}), \end{aligned} \tag{9}$$

where $h : \mathbb{R}^n \to \mathbb{R}$. So long as our inputs are not all dimensionless, we have $n < m$ and Equation 9 is a *ridge runction*; an equation with low-dimensional structure which is constant along those directions orthogonal to the columns of $\boldsymbol{V}$.[12] We call the range $\mathcal{R}(\boldsymbol{V})$ the *Pi Subspace*. Note that h takes as inputs linear combinations of $\boldsymbol{x}$, defined by $\boldsymbol{\xi} = \boldsymbol{V}^T \boldsymbol{x}$. This functional form is very amenable to study via Active Subspaces. In fact, consider the $\boldsymbol{C}$ matrix for Equation 9

$$\begin{aligned} \boldsymbol{C} &= \int_{\mathbb{R}^m} (\nabla_{\boldsymbol{x}} h(\boldsymbol{V}^T \boldsymbol{x}))(\nabla_{\boldsymbol{x}} h(\boldsymbol{V}^T \boldsymbol{x}))^T \rho(\boldsymbol{x})\, d\boldsymbol{x}, \\ &= \boldsymbol{V} \int_{\mathbb{R}^m} (\nabla_{\boldsymbol{\xi}} h(\boldsymbol{V}^T \boldsymbol{x}))(\nabla_{\boldsymbol{\xi}} h(\boldsymbol{V}^T \boldsymbol{x}))^T \rho(\boldsymbol{x})\, d\boldsymbol{x} \boldsymbol{V}^T, \\ &= \boldsymbol{V} \boldsymbol{T} \boldsymbol{V}^T, \end{aligned} \tag{10}$$

where $\nabla_{\boldsymbol{x}}$ is the gradient with respect to $\boldsymbol{x}$, and so on. Thus the Active Subspace of π is a subspace of the Pi Subspace. This Pi Subspace contains the powers associated with the physical inputs used to form dimensionless parameters. Thus the Active Subspace, being a subspace of $\mathcal{R}(\boldsymbol{V})$, inherits the same interpretation; its elements define products of the inputs, which are dimensionless parameters. We can make this explicit by exponentiating the Active Variables, which yields

$$\begin{aligned} \exp(\boldsymbol{y}) &= \exp(\boldsymbol{W}_1^T \boldsymbol{x}), \\ &= \exp(\boldsymbol{W}_1^T \log(\boldsymbol{q})), \\ &= \begin{pmatrix} \pi_1' \\ \vdots \\ \pi_k' \end{pmatrix}, \end{aligned} \tag{11}$$

thus the Active Variables have a 1-to-1 correspondence with dimensionless parameters, defined by the Active Subspace vectors. For this reason, we call the π_i' *active dimensionless parameters*. At this point,

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

there are a few important obervations to note: First, the π_i' are not necessarily the same as those which follow from classical dimensional analysis. The π_i which follow from the Buckingham π theorem are not unique, in the sense that they can be recombined as products to define an equally valid basis of dimensionless parameters. Rather, the π_i' which follow from the Active Subspace procedure above are *ranked in terms of their contributions to averaged changes in our QoI*. The $\boldsymbol{C}$ matrix is computed for a given system with defined parameter bounds; if its eigenvalues are distinct, this matrix has unique eigenvectors. The π_i' computed from $\boldsymbol{C}$ inherit this uniqueness property.[b]

Second, note that the dimension of the Active Subspace k is not necessarily the same as that of the Pi Subspace n. Note that Equation 10 implies that the Active Subspace is a contained by the Pi Subspace; that is, $\mathcal{R}(\boldsymbol{W}_1) \subseteq \mathcal{R}(\boldsymbol{V})$. Thus, we have $k \leq n$. The Buckingham π theorem states how many dimensionless parameters it is possible to form in a given system, but provides no means to determine which affect the QoI and which do not. If $k < n$ strictly, then we have identified those dimensionless parameters which appreciably affect the output, and can treat with less attention those which do not. This observation is at the core of design insight – a deep connection between the physics of the problem, the variables controlling the desired performance, and a coherent and computable mathematical structure.

These two properties provide an answer to an old question in Dimensional Analysis; that is, how to identify *relevant* dimensionless parameters. When provided with a model or data, the Active Subspace procedure provides the means to do just that. We will demonstrate these connections in the following example.

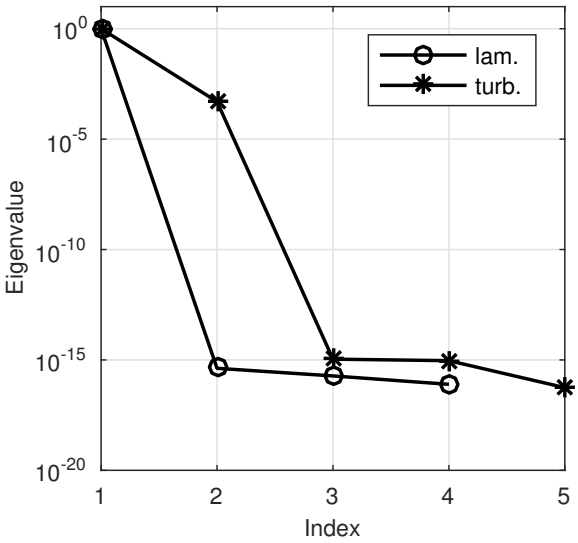

Figure 4. Normalized eigenvalues of $\boldsymbol{C}$ for pipe flow study. A zero eigenvalue indicates that no change occurs along the associated direction. Note that the eigenvalues here imply that the laminar case has a 1D Active Subspace, while the turbulent case is two dimensional. Note that since the leading eigenvalue in the turbulent case is several orders of magnitude larger than the second, the first dimensionless parameter is more important – on average – than the second.

D. Active Subspace of Pipe Bulk Velocity

We study the Active Subspace of the function in the form of Equation 9, thus we consider the equation in terms of log-transformed variables $\boldsymbol{x} = \log(\boldsymbol{q})$, which introduces some algebraic manipulations into the model. To compute the Active Subspace we define two operating scenarios, corresponding to laminar and turbulent conditions. These parameter bounds are defined in Appendix A, Tables 6 and 7 respectively. Note that we select a uniform weight function ρ indicating no specific preference among conditions within the input range.

To estimate the Active Subspace, we used finite differences to approximate gradients, and employed a tensor product Gauss-Legendre quadrature with 11 points (161051 points in five dimensions) in order to approximate the integrals of the $\boldsymbol{C}$ matrix. This was sufficient for 10 digits of accuracy in the eigenvalues. The eigenvalues resulting from this computation are shown in Figure 4, while the eigenvectors are summarized in Tables 2 and 3.

From the eigenvalues of $\boldsymbol{C}$ (Fig. 4), we conclude that the Active Subspace is 1-dimensional in the laminar case, and 2-dimensional in the turbulent case. This makes sense when compared against our model; Equation 4 relies on just one product of inputs, while Equation 5 depends on two. Studying the eigenvectors shows that the Active Subspace give us additional information.

Consider exponentiating the Active Variables in the laminar case, as in Equation 11. Since the Active Subspace is one-dimensional, we have just one Active Variable, defined by the leading eigenvector w_1 in Table

[b]Note that eigenvectors are defined up to a scalar multiple; propagating this property through exponentiation, we find that the resulting π_i' can be defined uniquely up to a power.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

	w_1	w_2	w_3	w_4	w_5
ρ	-0.2582	0.4946	-0.7873	0.2625	0
μ	0.5164	0.7361	0.1568	-0.4086	0
D	-0.7746	0.4019	0.4834	-0.0694	0
ϵ	0	0	0	0	-1.0000
$\frac{\Delta P}{L}$	-0.2582	-0.2281	-0.3492	-0.8714	0

Table 2. Eigenvectors of C in pipe flow experiment in the laminar case.

	w_1	w_2	w_3	w_4	w_5
ρ	-0.2408	0.1931	0.9477	-0.0796	0.0176
μ	0.4815	-0.3861	0.1382	-0.6569	0.4105
D	-0.8035	-0.2621	-0.1919	-0.4976	-0.0357
ϵ	0.0812	0.8412	-0.1919	-0.4976	-0.0357
$\frac{\Delta P}{L}$	-0.2408	0.1931	-0.0957	0.2587	0.9103

Table 3. Eigenvectors of C in pipe flow experiment in the turbulent case.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

2. Since eigenvectors are only unique up to scalar multiples, we normalize by the smallest vector element, for reasons that will become apparent momentarily. We find

$$\begin{aligned}\exp(y_1) &= \exp([1.0, -2.0, 3.0, 0.0, 1.0]\log(\boldsymbol{q})),\\ &= \rho^{1.0}\mu^{-2.0}D^{3.0}\epsilon^{0.0}\left(\frac{\Delta P}{L}\right)^{1.0},\\ &= \frac{\Delta P}{L}\frac{\rho D^3}{\mu^2}.\end{aligned} \tag{12}$$

Returning to a relation of the form of Equation 6, we may constrain the law governing Re, and re-dimensionalize to study our dimensional QoI

$$\begin{aligned}Re &= f(\frac{\Delta P}{L}\frac{\rho D^3}{\mu^2}),\\ V &= \frac{\mu}{\rho D}f(\frac{\Delta P}{L}\frac{\rho D^3}{\mu^2}),\end{aligned} \tag{13}$$

where f remains an unknown function, to be determined through some other process.

Note that Buckingham π suggested that up to two dimensionless parameters governed our QoI; our analysis has found that ϵ does not affect V in the laminar parameter range, and so there is only one active dimensionless parameter. Thus the Active Subspace method is *able to find the governing dimensionless parameter algorithmically!* Note that this is a substantial improvement over the results from Classical Dimensional Analysis; rather than concluding that Re depends on two non-unique dimensionless parameters, we can determine that in the laminar regime, Re depends only on the dimensionless parameter given in Equation 12. Of course, this comparison against Classical Dimensional Analysis is somewhat unfair – the Active Subspace analysis was augmented with a great deal of data. However, it is important to note that this analysis was carried out *algorithmically*; rather than scrutinizing plots of data and employing the 'method of concentrated staring',[13] we arrive at our conclusion in an automated fashion.[c]

The results in the turbulent case are interpreted in much the same way; the eigenvectors define dimensionless parameters, just as in Equation 12. In this case, since the Active Subspace is 2-dimensional there are two active dimensionless parameters. Here we introduce a simple technque to aid in analyzing the results of an Active Subspace analysis. In the laminar case, the powers rounded to simple integer values; however,

[c]Here we would like to recognize the wit of Prof. S. Shvartsman, who coined the wonderfully descriptive phrase used above.

there is no guarantee that such a simplification is possible, which can make interpreting the resulting dimensionless parameters challenging. However, we may choose to represent our results in a basis of pre-defined dimensionless parameters. This can be particularly helpful for interpreting our results in the context of classical dimensionless parameters. Since the Active Subspace directions live in the nullspace of $\boldsymbol{D}$, we may uniquely represent these vectors in terms of a basis for $N(\boldsymbol{D})$, which follows from the solution of a simple linear system.

To illustrate, select $\frac{\epsilon}{D} \equiv \pi_1$ and $\frac{\Delta P}{L}\frac{\rho D^3}{\mu^2} \equiv \pi_2$ to be our basis of dimensionless parameters, and let $\boldsymbol{W}_1$ be the output of our Active Subspace computation. The basis above corresponds to the matrix

$$\boldsymbol{B}^T = \begin{bmatrix} 1 & -2 & 3 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 \end{bmatrix}. \tag{14}$$

The powers defining our active dimensionless parameters are given by the solution to $\boldsymbol{BP} = \boldsymbol{W}_1$. The presentation of Equation 15 follows from the computation detailed above.

$$\begin{aligned} \pi_1' \equiv \exp(y_1) &= \rho^{-0.2408}\mu^{0.4815}D^{-0.8035}\epsilon^{0.0812}\left(\frac{\Delta P}{L}\right)^{-0.2408}, \\ &= \left(\frac{\epsilon}{D}\right)^{0.0812}\left(\frac{\Delta P}{L}\frac{\rho D^3}{\mu^2}\right)^{-0.2408}, \\ \pi_2' \equiv \exp(y_2) &= \rho^{0.1931}\mu^{-0.3861}D^{-0.2621}\epsilon^{0.8412}\left(\frac{\Delta P}{L}\right)^{0.1931}, \\ &= \left(\frac{\epsilon}{D}\right)^{0.8412}\left(\frac{\Delta P}{L}\frac{\rho D^3}{\mu^2}\right)^{0.1931}. \end{aligned} \tag{15}$$

Thus, we find

$$\begin{aligned} Re &= f(\pi_1', \pi_2'), \\ V &= \frac{\mu}{\rho D} f(\pi_1', \pi_2'), \end{aligned} \tag{16}$$

Strictly speaking, the Active Subspace in this case is 2-dimensional, so we may not eliminate any dimensionless parameters entirely. However, the eigenvalues of $\boldsymbol{C}$ give us additional information; since the first eigenvalue is several orders of magnitude larger than the trailing eigenvalues, this tells us that the QoI changes on average more along the first direction. Thus, the first Active Variable π_1' accounts for most of the variability in the function. In the context of design, this ranking of the Active Variables allows us to focus our attention on the inputs which change our QoI the most.

E. Dimensional Analysis: Insights

In this section, we have illustrated a connection between Dimensional Analysis and Active Subspaces, which leads to deep insights about the physical system. By computing the Active Subspace on the log-transformed version of the problem (Eq. 9), we can identify *relevant* dimensionless parameters, providing additional information over Classical Dimensional Analysis. From an engineering design perspective, this insight allows a designer to focus on relevant inputs to a system, and provides a meaningful interpretation of the Active Variables in terms of familiar dimensionless parameters.

The pipe flow example illustrates how this technique might be used in practice. By computing the Active Subspace on the log-transformed problem, we determined that our dimensionless QoI depends only on the dimensionless pressure in the laminar regime, and on combinations of dimensionless pressure and relative roughness in the turbulent regime. Note that in the turbulent case, we are also able to rank the governing dimensionless parameters in terms of their importance to our QoI. It is also important to note that these insights are the product of an automated process – rather than studying numerous summary plots, the Active Subspace procedure gives us the active dimensionless parameters algorithmically.

The process above works when our mathematical description of the problem is 'complete', in the sense that we consider all variables which are important to our QoI. In the case where we are missing one or more important variables, the resulting Active Subspace directions do not necessarily describe dimensionless

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

parameters. We make this failure mode clear in the following section, and provide a technique to detect such *hidden parameters*.

IV. Hidden Parameters and Active Subspaces

In this section, we consider a failure mode of Classical Dimensional Analysis – the case where one or more important parameters have been omitted from the analysis. We begin with a mathematical formulation of this phenomenon and present a technique which, in some cases, allows for the detection of these *hidden parameters*. We then demonstrate this technique on the pipe flow problem described previously.

A. Formulation

The Buckingham π theorem applies to physical laws due to their property of *dimensional homogeneity*; that is, the property where both sides of a relation have the same units. Relations which are not physical and yet still dimensionally homogeneous enjoy the same dependence on dimensionless parameters, while conversely the Buckingham π theorem does not apply to relations which are not dimensionally homogeneous. However, the latter case can be somewhat obscured, depending on our choice of input variables. If the true physical relation depends on the inputs $\{q_1, \ldots, q_m\}$, but we choose a subset of these (e.g. as a result of assumptions), the resulting interrelation of the chosen $\boldsymbol{q}$ and π may appear to be dimensionally inconsistent. We will make this case more apparent with the following example.

Consider our physical pipe flow problem from above. The diagram above (Fig. 3) shows that the dependence in the system on ϵ only exists above the critical Reynolds number; all the behavior of the system is captured in the laminar regime by considering the subset of parameters $\{\rho, \mu, D, \frac{\Delta P}{L}\}$. Dimensional Analysis in this case will predict only one dimensionless parameter. However, if we move to the turbulent regime without reintroducing ϵ, we will miss important behavior. Most importantly, the incorrect conclusion of just one Dimensionless Parameter will prevent the collapse of data, which will confound our analysis.

However, we may use the local properties of the ridge function form of our physical law to help detect this scenario. Suppose we are considering the log variables $\boldsymbol{x}$, and there exists some true physical law as in Equation 9, but we have chosen (whether knowingly or not) a subset of parameters $\boldsymbol{x}_e = \boldsymbol{E}_e^T \boldsymbol{x}$. The vector $\boldsymbol{x}_e \in \mathbb{R}^p$ is the vector of our p *exposed parameters* from the full set of m inputs, while the *adapter matrix* $\boldsymbol{E}_e \in \mathbb{R}^{m \times p}$ extracts the appropriate elements from an m-vector (its rows are standard basis vectors chosen to accomplish this). Note that left multiplication by $\boldsymbol{E}_e^T$ deletes rows, while left multiplication by $\boldsymbol{E}_e$ pads the input with zeros. Similarly, we define the *hidden parameters* $\boldsymbol{x}_h = \boldsymbol{H}_h^T \boldsymbol{x} \in \mathbb{R}^{m-p}$, such that

$$\boldsymbol{x} = \boldsymbol{E}_e \boldsymbol{x}_e + \boldsymbol{E}_h \boldsymbol{x}_h. \tag{17}$$

In practice, we vary only the exposed parameters $\boldsymbol{x}_e$, while the hidden parameters $\boldsymbol{x}_h$ adopt some nominal values. We can rewrite Equation 9 as $\pi = h(\boldsymbol{V}^T \boldsymbol{E}_e \boldsymbol{x}_e + \boldsymbol{V}^T \boldsymbol{E}_h \boldsymbol{x}_h)$. Define $\boldsymbol{V}_e = \boldsymbol{E}_e^T \boldsymbol{V}$ and $\boldsymbol{V}_h = \boldsymbol{E}_h^T \boldsymbol{V}$, and compute the gradient with respect to our exposed and hidden parameters

$$\begin{aligned} \nabla_{\boldsymbol{x}_e} \pi &= \boldsymbol{V}_e \nabla_{\boldsymbol{\xi}} h(\boldsymbol{\xi}), \\ \nabla_{\boldsymbol{x}_h} \pi &= \boldsymbol{V}_h \nabla_{\boldsymbol{\xi}} h(\boldsymbol{\xi}), \end{aligned} \tag{18}$$

where $\boldsymbol{\xi} = \boldsymbol{V}^T \boldsymbol{x}$ as before. By applying chain rule to the full gradient $\nabla_{\boldsymbol{x}} \pi$, we find

$$\boldsymbol{V} \nabla_{\boldsymbol{\xi}} h(\boldsymbol{\xi}) = \boldsymbol{E}_e \boldsymbol{V}_e \nabla_{\boldsymbol{\xi}} h(\boldsymbol{\xi}) + \boldsymbol{E}_h \boldsymbol{V}_h \nabla_{\boldsymbol{\xi}} h(\boldsymbol{\xi}). \tag{19}$$

However, we know that $\boldsymbol{V}$ is a basis for the nullspace of $\boldsymbol{D}$; multiplying by $\boldsymbol{D}$ and substituting expressions from Equations 18, we find

$$\begin{aligned} \boldsymbol{D}\boldsymbol{E}_e \boldsymbol{V}_e \nabla_{\boldsymbol{\xi}} h(\boldsymbol{\xi}) &= -\boldsymbol{D}\boldsymbol{E}_h \boldsymbol{V}_h \nabla_{\boldsymbol{\xi}} h(\boldsymbol{\xi}), \\ \boldsymbol{D}_e \nabla_{\boldsymbol{x}_e} \pi &= -\boldsymbol{D}_h \nabla_{\boldsymbol{x}_h} \pi, \end{aligned} \tag{20}$$

where $\boldsymbol{D}_e = \boldsymbol{D}\boldsymbol{E}_e$ and $\boldsymbol{D}_h = \boldsymbol{D}\boldsymbol{E}_h$ are Dimension Matrices with the appropriate columns removed. Thus, we see that the fundamental dimensions of hidden parameters appear even as we compute gradients with respect to the exposed parameters alone, as made precise in Equation 20. In practice, if $\boldsymbol{D}_e \nabla_{\boldsymbol{x}_e} \pi \neq 0$, we know that the set of hidden parameters is non-empty. This can serve as a diagnostic tool to detect when

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

we are missing information. Note that Equation 20 gives us much more than binary information, though; the fundamental dimensions which appear in our analysis $\boldsymbol{D}_e \nabla_{\boldsymbol{x}_e} \pi$ are colinear with those of the hidden parameters $\boldsymbol{D}_h \nabla_{\boldsymbol{x}_h} \pi$. For instance, if a length scale appears in $\boldsymbol{D}_e \nabla_{\boldsymbol{x}_e} \pi$, we would know that a hidden parameter exists which has the units of L^r. Unfortunately, we cannot distinguish between powers; in the previous example, we cannot distinguish between a length, an area, a volume, etc. However, this information does provide clear information about existence, and some rough information about the character of the hidden parameters.[d]

As the Active Subspace of π with respect to the full parameters $\boldsymbol{x}$ is a subset of the Pi Subspace $\mathcal{R}(\boldsymbol{V})$, we also have that the Active Subspace with respect to the exposed parameters $\boldsymbol{x}_e$ is a subset of a reduced Pi Subspace $\mathcal{R}(\boldsymbol{V}_e)$. Thus, we can choose to check individual gradient samples, or consider an aggregated measure by studying the Active Subspace. We demonstrate using the Active Subspace to detect hidden parameters in the example below.

B. Hidden Parameter Detection Example

Consider the pipe flow example described above; we consider exposed parameters $\{\rho, \mu, D, \frac{\Delta P}{L}\}$ and hidden parameter ϵ. In practice, this means we vary the exposed parameters, and set ϵ equal to a nominal value. This yields the adapter matrices

$$\boldsymbol{E}_e = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \boldsymbol{E}_h = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \tag{21}$$

and the Dimension Matrix for the exposed parameters

$$\boldsymbol{D}_e = \begin{bmatrix} 1 & 1 & 0 & 1 \\ -3 & -1 & 1 & -2 \\ 0 & -1 & 0 & -2 \end{bmatrix}. \tag{22}$$

We approximate $\boldsymbol{C}$ via Monte Carlo sampling ($n = 100$) in the laminar and turbulent parameter regimes, and consider the product $\boldsymbol{D}_e \boldsymbol{W}_1$ in the two cases, shown in Tables 4 and 5. Note that in the turbulent case, the results suggest that a hidden parameter exists, and provides information that suggests a missing length scale.

	$\boldsymbol{D}_e \boldsymbol{w}_1$	$\boldsymbol{D}_e \boldsymbol{w}_2$	$\boldsymbol{D}_e \boldsymbol{w}_3$	$\boldsymbol{D}_e \boldsymbol{w}_4$
M	-0.0000	-1.2447	-1.0279	0.6278
L	0.0000	3.5057	0.6750	-1.5014
T	0.0000	0.4242	1.3452	-1.7351

Table 4. Product $\boldsymbol{D}_e \boldsymbol{W}_1$ for the laminar case. The eigenvalues of $\boldsymbol{C}$ suggest a one-dimensional Active Subspace, and the leading term is dimensionless (the zero vector), therefore we do not suspect any hidden parameters are present.

[d] There is another important failure mode in this process, which is when an *entire* dimensionless parameter is missing. In this case, the missing fundamental dimensions may cancel, leaving the parameter entirely hidden.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

	$\boldsymbol{D_e w_1}$	$\boldsymbol{D_e w_2}$	$\boldsymbol{D_e w_3}$	$\boldsymbol{D_e w_4}$
M	-0.0000	-0.0000	1.1469	1.2979
L	-0.0887	-1.5786	-2.8237	-2.1277
T	-0.0000	0.0000	-0.0871	-2.2344

Table 5. Product $\boldsymbol{D_e W_1}$ for the turbulent case. The eigenvalues of $\boldsymbol{C}$ suggest a two-dimensional Active Subspace, and the leading terms are not dimensionless, therefore we suspect hidden parameters are present. Additionally, the leading vectors have a nonzero entry in length only, suggesting the hidden parameter is some form of lengthscale, which is consistent with having hidden ϵ.

C. Hidden Parameters: Insights

In this section, we have mathematically formulated the hidden parameter problem and proposed a technique to detect such hidden parameters. In the case where we are missing an important dimensional parameter, as in hiding ϵ previously, we may employ the aforementioned analysis to detect hidden parameters, and provide some information about their dimensional content. This allows a designer to study 'black box' codes, and detect cases where upstream assumptions have been made by the implementer, and certain parameters have been hidden from the user.

The insights provided by a hidden parameter analysis are different from the others presented in this work; whereas active dimensionless parameters and active deformations (Sec. V) provide exploitable structure in a design problem, the presence of hidden parameters suggests the possibility [e] that *something in the analysis is incorrect*. However, this insight is no less valuable! Determining code correctness is crucial in Verification and Validation,[14] and neglecting important variables in an analysis may prevent an optimal design from being reached.

One may also regard the hidden parameter study as a means to test insights; if we have determined that a particular set of variables are important for our QoI in a particular parameter range, we can test whether these remain the only relevant variables in a different parameter regime. This perspective is illustrated in the aforementioned hidden parameter study.

Our examples so far have considered a physical system which is 'rich' in dimensional information; the input parameters largely have different fundamental dimensions, and so their combinations are constrained by dimensional homogeneity. In the next section, we consider a case where all inputs have the same fundamental dimensions, in which case we lose much of the power of Buckingham π.

V. Shape Design and Active Subspaces

In this section we consider the general problem of shape design, and draw connections between geometric deformations and Active Subspaces to identify *active deformations*. We consider an example problem of modifying an airfoil to best vary the lift coefficient. We begin with the formulation of our design problem, then draw connections between Active Subspaces and relevant deformations of the baseline geometry, and present results from a specific design problem. Finally, we connect these insights to classical aerodynamics results, and suggest the analysis outlined as a general tool for generating novel design insights.

A. Formulation of the Design Problem

We consider a baseline geometry defined by the NACA0012 profile, a symmetric airfoil common in aerodynamic studies.[15] We vary the Mach number in our studies, while keeping fixed a modest angle of attack (AoA) ($\alpha = 1.25°$) and subsonic inlet conditions ($M = 0.8$), with standard temperature and pressure conditions.

To parameterize the airfoil design, we employ a Free-Form Deformation (FFD) strategy.[16] In this approach, a bounding box is defined around our airfoil, and parameterized as a Bézier solid. A set of control points $P_{i,j,k}$ are defined on the box, which is itself parameterized by

$$X(u,v,w) = \sum_{i=0}^{l} \sum_{j=0}^{m} \sum_{k=0}^{n} P_{i,j,k} B_i^l(u) B_j^m(v) B_k^n(w), \tag{23}$$

[e] Note that there is occasionally good reason to hide a parameter value which is truly fixed; e.g. the gas constant

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

where l, m, n are the degrees of the FFD function, $u, v, w \in [0, 1]$ are the parameteric coordinates, and B_i^l, B_j^m, B_k^n are Bernstein polynomials.[f] The coordinates of the airfoil are mapped to the Bézier box, and once deformation has been applied the coordinates of the new airfoil are found by via Equation 23. In our application, we parameterize our FFD box with 42 control points (half along the top, half along the bottom), and vary only the vertical displacement of the internal points for a total of 38 design variables. The selection of internal points is done to ensure the AoA remains constant; otherwise, the optimizer could bend the tips of the airfoil to change the effective AoA.

Note that the coordinates $P_{i,j,k}$ are our input parameters, which all have units of length. Thus, we know that an analysis via Buckingham π will be uninformative; given a set of m lengths, we will find that up to $m - 1$ dimensionless parameters are relevant. While it is possible to follow the procedure outlined above to identify relevant dimensionless parameters, we instead adopt a graphical interpretation of the Active Subspace as *active deformations*. Since the design variables define deformations of the airfoil, directions in parameter space define coherent shape deformations. Additionally, since the Active Subspace ranks these directions in terms of their importance to our QoI, we may organize these deformations to determine which are relevant to our interests. This interpretation is demonstrated in the following sections.

In all cases below, we consider the lift coefficient as our QoI. We solve the Euler equations with the open-source CFD code SU2, using an adjoint solver to compute sensitivities of our QoI with respect to the design variables.[16] A representative flow solution is given in Appendix B. Due to the relative high dimensionality of the design problem, we estimate the Active Subspace via Monte Carlo sampling, and check the quality of our estimate via a bootstrap analysis.

B. Active Subspace of the Lift Coefficient

We define reasonable hypercube bounds on the design variables, and approximate Equation 1 by an estimated $\hat{\boldsymbol{C}}$ matrix computed via Monte Carlo sampling ($n = 10^3$). To check the quality of our Active Subspace estimator, we perform bootstrap resampling on our gradients ($b = 10^4$ full resamples of the gradients) and recompute the Active Subspace on each new draw. Figure 5 plots bootstrap intervals of the eigenvalues, which help determine the dimension and quality of our Active Subspace estimate. The same figure plots bootstrap intervals for the subspace distance, which characterize the stability of the Active Subspace estimate.[5]

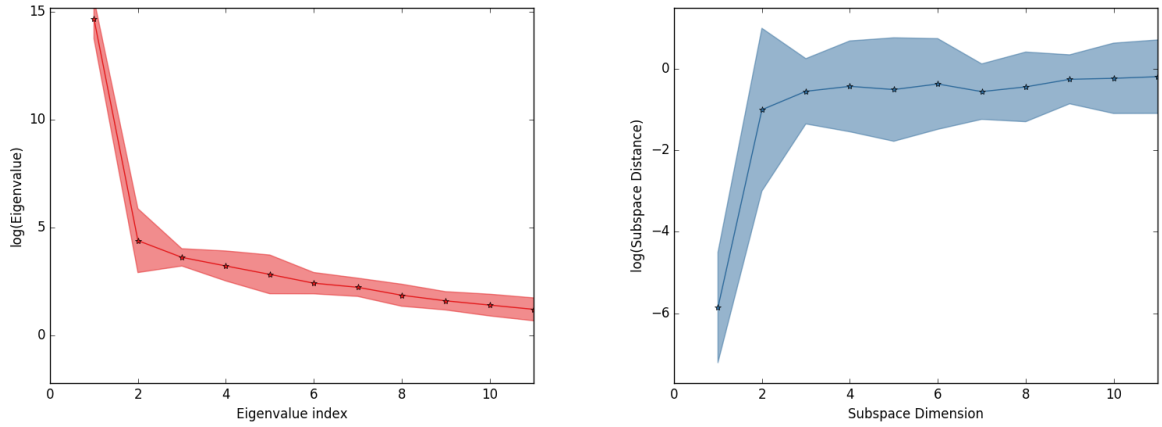

Figure 5. Bootstrap intervals of the eigenvalues of $\hat{\boldsymbol{C}}$ and subspace distance for NACA0012 test case. The left image shows bootstrap intervals of the eigenvalues, while the right displays those of the subspace distance. Stars denote the estimates derived from the entire sample set, while curves and shaded bounds show the bootstrap mean and 95% confidence intervals, respectively. The dimension of the Active Subspace is chosen by seeking eigenvalue gaps; thus, the spectrum of $\hat{\boldsymbol{C}}$ suggests a one-dimensional Active Subspace. The subspace distance is computed between the Active Subspace estimate $\hat{\boldsymbol{W}}_1$ and a bootstrap estimate of the same quantity. This quantity is bounded above by 1; the smaller this value, the more stable our estimate $\hat{\boldsymbol{W}}_1$.

The spectrum presented in Figure 5 suggests that a one-dimensional Active Subspace accounts for the vast majority of variability in our QoI; thus, we consider only the leading eigenvector $\hat{\boldsymbol{w}}_1$ of $\hat{\boldsymbol{C}}$. To visualize

[f] Note that for our application, we need only deformation within two dimensions. We define an FFD box about our airfoil with depth in a superfluous z-dimension for technical reasons, but vary the control points to achieve 2-dimensional deformation.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

the deformation associated with this leading eigenvector, we fix the design variables at a point $\boldsymbol{x} = 0 + \alpha\hat{\boldsymbol{w}}_1$, where α is small, and on the order of the bounds of our design variables. The results of this deformation along $\hat{\boldsymbol{w}}_1$ are shown in Figure 6.

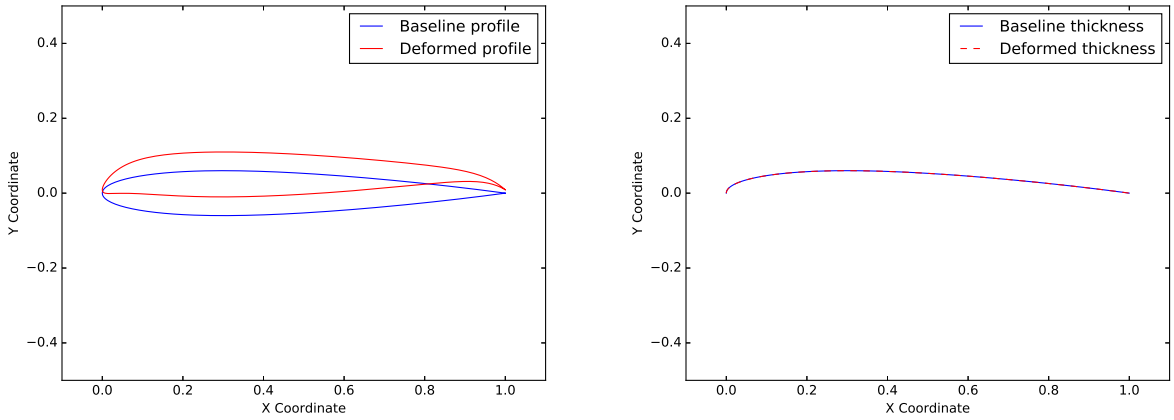

Figure 6. The left image compares the baseline geometry against the airfoil deformed along the leading Active Subspace direction. This deformation corresponds to the camber line, as evidenced by the thickness curves plotted in the right image. Since an airfoil may be uniquely decomposed into a linear combination of its thickness and camber, and since both profiles have the same thickness curve, we conclude the deformation is modifying the camber of the airfoil alone. Taking the L1 norm of the difference between thickness curves yields 5.6948×10^{-7}**.**

Note that the results in Figure 6 imply that *the active deformation is the camber line*! That is, deforming along the leading eigenvector $\hat{\boldsymbol{w}}_1$ is equivalent to increasing/decreasing the camber of the airfoil. Thus, the Active Subspace procedure has identified a fundamental aerodynamic concept algorithmically.

C. Shape Design: Insights

Through this example, we have seen how we may interpret the Active Subspace in the context of shape design. The Active Directions define active deformations of our baseline geometry, and their associated eigenvalues provide a means to rank the plethora of deformations available to us as designers. Inviscid thin airfoil theory predicts that among the shape deformations one could make, only the camber line affects the lift coefficient – this fact is discovered algorithmically by the Active Subspace. While above we recover classical results, the truly exciting potential of this viewpoint is the potential to develop novel insights in different areas.

These insights are useful both from a computational and conceptual standpoint. Usually, Active Subspaces are used as a dimension reduction technique, in order to make high-dimensional problems tractable. By interpreting the Active Subspace directions as shape deformations, one may build a conceptual understanding which has value beyond the inherent dimension reduction. For example, consider the multi-component wing system noted in the Introduction. The insight of the camber line guides the design of such systems, which modify the effective camber of the airfoil during operation, in order to achieve different performance characteristics. While the historical insight of the importance of camber to lift was the hard-won product of classical analysis, our example above shows that the Active Subspace procedure can discover this insight *algorithmically*. The application of this technique to new design areas has the potential to reveal novel insights of a similar caliber, and to serve as a powerful conceptual technique for engineering designers.

VI. Summary & Conclusions

In this work we have considered using the Active Subspace to develop design insight; that is, an exploitable, conceptual understanding of the engineering problem at hand. In the context of a 'dimensionally rich' setting where Dimensional Analysis may be employed, we have described how to use the Active Subspace to identify active dimensionless parameters, and to detect when hidden parameters are missing from our analysis. We demonstrated this approach in a pipe flow example, by identifying the active dimension-

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

less parameters in the laminar and turbulent regimes, and by providing an example of hidden parameter detection. In the context of design via shape deformation, we presented a graphical interpretation of the Active Subspace as active deformations, which may be leveraged both in a computational setting, and in conceptual design. We demonstrated this approach in the context of designing an airfoil with a focus on the lift coefficient. The resulting active deformation was found to correspond to the camber line, which is in agreement with classical thin airfoil theory.

The primary contribution of this work is an elucidation of the connection between the Buckingham π Theorem and Active Subspaces, and a practical recipe for addressing long-standing issues in Dimensional Analysis; namely, the identification of relevant (active) dimensionless parameters, and the detection of important missing variables – hidden parameters.

The other major contribution of this work is a general process for developing conceptual understanding in shape design problems. The example above demonstrated the discovery of the dependence of lift on the camber line *algorithmically*. While this recovers classical results in our simple example, the technique has the potential to discover similar insights in novel problems, which may prove a boon to engineering designers searching for novel insights.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

VII. Acknowledgements

This material is based on work supported by Department of Defense, Defense Advanced Research Project Agency′s program Enabling Quantification of Uncertainty in Physical Systems. The first author′s work is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-114747.

References

[1]Pahl, G. and Beitz, W., *Engineering Design: A Systematic Approach*, Springer, 2nd ed., 1996.

[2]Anderson, J. J., *Fundamentals of Aerodynamics*, McGraw-Hill, 3rd ed., 2001.

[3]Smith, A., "High-Lift Aerodynamics," *AIAA 6th Aircraft Design, Flight Test and Operations Meeting*, Vol. 12, No. 6, 1974.

[4]Liepmann, H. and Roshko, A., *Elements of Gasdynamics*, Dover Publications, 1st ed., 1957.

[5]Constantine, P., *Active Subspaces: Emerging Ideas for Dimension Reduction in Parameter Studies*, SIAM Philadelphia, 2015.

[6]T. Lukaczyk, F. Palacios, J. A. and Constantine, P., "Active subspaces for shape optimization," *10th AIAA Multidisciplinary Design Optimization Conference*, 2014.

[7]Constantine, P., Emory, M., Larsson, J., and Iaccarino, G., "Exploiting active subspaces to quantify uncertainty in the numerical simulation of the HyShot II scramjet," *Journal of Computational Physics*, Vol. 302, 2015, pp. 1 – 20.

[8]Jefferson, J. L., Gilbert, J. M., Constantine, P. G., and Maxwell, R. M., "Active subspaces for sensitivity analysis and dimension reduction of an integrated hydrologic model," *Computers & Geosciences*, Vol. 83, 2015, pp. 127 – 138.

[9]White, F. M., *Fluid Mechanics*, McGraw-Hill, New York, 2011.

[10]Colebrook, C. F. and White, C. M., "Experiments with fluid friction in roughened pipes," *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, Vol. 161, No. 906, 1937, pp. 367–381.

[11]Buckingham, E., "On Physically Similar Systems; Illustrations of the Use of Dimensional Equations," *Phys. Rev.*, Vol. 4, Oct 1914, pp. 345–376.

[12]Constantine, P. G., del Rosario, Z., and Iaccarino, G., "Many physical laws are ridge functions," *ArXiv e-prints*, May 2016.

[13]Shvartsman, S., personal communication.

[14]Oberkampf, W. and Roy, C., *Verification and validation in scientific computing*, Cambridge University Press, 2010.

[15]Jacobs, E. N., Ward, K. E., and Pinkerton, R. M., "The characteristics of 78 related airfoil sections from tests in the variable-density wind tunnel," 1933.

[16]Palacios, F., Colonno, M. R., Aranake, A. C., Campos, A., Copeland, S. R., Economon, T. D., Lonkar, A. K., Lukaczyk, T. W., Taylor, T. W. R., , and Alonso, J. J., "Stanford University Unstructured (SU2): An open-source integrated computational environment for multi-physics simulation and design," *AIAA Aerospace Sciences Meeting and Exhibit*, 2013.

VIII. Appendix

A. Parameter bounds for pipe flow example

The following tables define parameter bounds for the pipe flow example in Section III.

fluid density	ρ	1.0×10^{-1}	1.4×10^{-1}	$\mathrm{kg/m^3}$
fluid viscosity	μ	1.0×10^{-6}	1.0×10^{-5}	$\mathrm{kg/(ms)}$
pipe diameter	D	1.0×10^{-1}	$1.0 \times 10^{+0}$	m
pipe roughness	ε	1.0×10^{-3}	1.0×10^{-1}	m
pressure gradient	$\frac{\Delta P}{L}$	1.0×10^{-7}	1.0×10^{-6}	$\mathrm{kg/(ms)^2}$

Table 6. Parameter bounds for the laminar flow case.

fluid density	ρ	1.0×10^{-1}	1.4×10^{-1}	$\mathrm{kg/m^3}$
fluid viscosity	μ	1.0×10^{-6}	1.0×10^{-5}	$\mathrm{kg/(ms)}$
pipe diameter	D	1.0×10^{-1}	$1.0 \times 10^{+0}$	m
pipe roughness	ε	1.0×10^{-3}	1.0×10^{-1}	m
pressure gradient	$\frac{\Delta P}{L}$	1.0×10^{-1}	$1.0 \times 10^{+1}$	$\mathrm{kg/(ms)^2}$

Table 7. Parameter bounds for the turbulent flow case.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090

B. Representative Flow Solution

The following figure gives a representative flow solution for the NACA0012 design case.

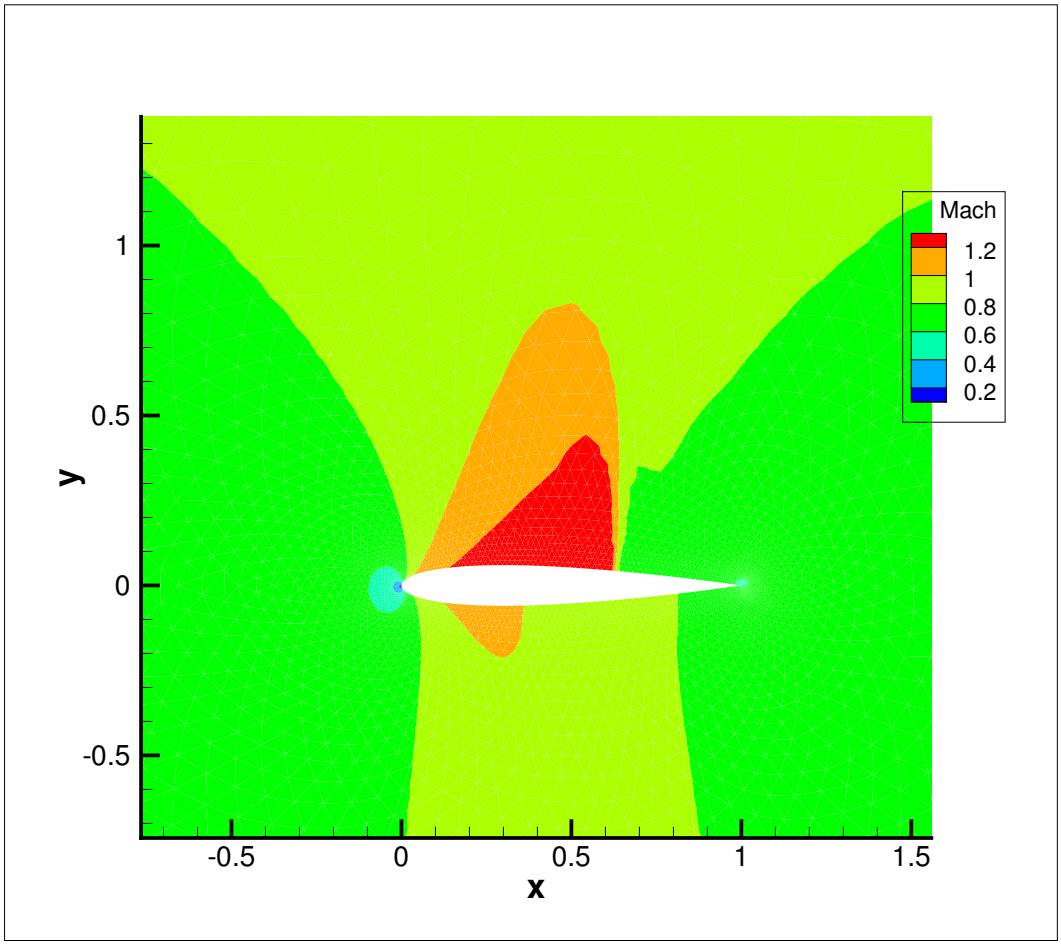

Figure 7. Inlet condition $M_\infty = 0.8$; Mach contours over a NACA0012 airfoil.

Downloaded by STANFORD UNIVERSITY on August 20, 2018 | http://arc.aiaa.org | DOI: 10.2514/6.2017-1090