Data-driven Stochastic Optimization on Manifolds for Additive Manufacturing

Myrl G. Marmarelis*, Roger G. Ghanem

Viterbi School of Engineering, University of Southern California

April 2020

Highlights

- We derive a mathematical formulation to infer the dependency structure between build parameters in laser powder-bed fusion and measured tensile properties.
- The method employs a novel kernel-density estimation technique within a Riemannian manifold embedding that is approximated via diffusion maps.
- Monte Carlo sampling reveals the expected Pareto frontier of the three outputs, and allows for stochastic optimization of arbitrary objectives.

Abstract

Laser powder-bed fusion is an increasingly attractive modality for additive manufacturing. Simulating the precise behavior of the process is infeasible, leading researchers to seek proxy solutions in machine learning. We demonstrate a novel adaptation of a technique called diffusion maps to infer the dependence structure between build parameters and material properties of interest on an approximated Riemannian manifold, when attaining a sufficient number of samples is cost prohibitive. We perform stochastic optimization to learn the efficient frontier of multi-objective combinations as well as to locate the set of parameters (hatch width, laser speed, and power) that putatively minimize the chance of failure, herein defined as lower bounds on the material properties (ultimate strength, yield strength, and elongation). Out-of-sample validation is performed and the results confirm our model's efficacy in optimizing coupons, and perhaps complex geometries in the future.

Keywords

Additive manufacturing; Diffusion maps; Kernel density estimation; Stochastic optimization; Laser powder-bed fusion.

1 Introduction

In the growing field of additive manufacturing (AM), the complex phenomena that are harnessed for metal part fabrication can be modeled generally and approximately by first principles. There are several competing AM technologies available in industry. The state-of-the-art modality employed in this study is termed laser powder-bed fusion [1], wherein a laser beam is traced upon a flat bed of metal powder to selectively melt and rapidly solidify the regions of that layer corresponding to the sliced digital specification, the cross-sectional area of the specific layer being irradiated. Sophisticated AM techniques

*Corresponding author—contact via myrlgmar@gmail.com.

attained traction in industries like defense, aerospace, and healthcare—all of which may be aptly characterized as risk-averse in their products or otherwise stated, conservative in the adoption of new manufacturing technologies due to the high stakes involved in the associated applications. AM does not enjoy the legacy and accrued knowledge of more traditional manufacturing disciplines, presenting an obstacle to complete adoption. It is therefore essential to gain a wholistic view of the intricate landscape of variables that affect these methods of fabrication. This work seeks to develop a methodology to predict the key process parameters in the modality of laser powder-bed fusion that yield optimized mechanical properties. We propose a technique to explore the tradeoffs between desirable yet conflicting outcomes, such as high elongation and high yield strength. We also demonstrate a probabilistic optimization scheme that minimizes the overall chance of a generalized failure event.

Process maps [2] outline the causal dependency of key quantities of interest to the relevant parameters of the build (e.g. laser power, laser scan speed, and hatch distance, where hatch distance is the distance between individual laser lines and is usually on the order of the melt-pool width). Recent published work pioneered the successful use of machine learning in estimating data-driven process maps [3] and in particular mitigating the chance of failed builds [4]. These endeavors used Support Vector Machines and Gaussian Mixture Models, respectively. Widely celebrated for their simplicity and adaptability, these methodologies were apt choices for a field as fresh to machine learning as additive manufacturing. Sophisticated schemes for automated experimental design within the context of the aforementioned process maps have also been developed [5], incorporating the feedback from subsequent results to inform the upcoming experiments.

Machine learning has emerged in the domain of materials science to approximate phenomena, the inner workings of which are still out of our grasp. The past couple of years have seen a movement to understand the propagation of fractures in mainly brittle materials via harnessing reduced-order models [6] or popular data-driven algorithms. Researchers attained success in developing proxy models to computationally expensive numerical simulators for increased efficiency without significant compromise to predictive accuracy [7]. Some implemented a deep-learning protocol and satisfied its thirst for data by augmenting the training sample with the model's own past predictions [8]. A similar problem was solved differently by Soize and Ghanem when they generated artificial samples with a stochastic process along an estimated manifold [9]; this approach served as inspiration for the present work. Our probabilistic formulation differs from the efforts of Rovinelli et al. [10] in that we discover implicit structure in the data using diffusion maps, whereas they postulate a specific analytic expression.

Laser powder-bed fusion, also known as selective laser melting and closely related to direct laser deposition, yields metal structures that are governed by the thermal intrinsics of the melt pools [11] traced by the laser. Many uncertain elements are due to the impurities in the process inputs (e.g. powder distribution), and agglomerated by a complex nonlinear interplay with the handful of build parameters. The need for a data-driven endeavor to serve as an efficient proxy for finite-element simulations, as well as obviate the excessive amount of costly experimentation in understanding the melt-pool behavior was addressed lately by means of Gaussian process regression [12]. Our work stands in contrast by two counts: first, our approach involves kernel density estimation on an approximate manifold embedding; second, we seek to attain the direct process maps, i.e. the outputs like tensile strength that validate the final product (as opposed to, and bypassing an intermediate build variable) as a function of the inputs, tailored to a particular configuration. The presented case study confirms our data-driven technique on cylindrical build coupons. It is effortlessly extendable to more complex geometries as desired: the model itself remains agnostic. We encourage a *ceteris paribus* experimental approach, in holding fixed all other controllable parameters not fed into the model. To investigate a cohort with different geometries would require to codify the variations concisely into a couple of real-valued dimensions in the parameter space. Exogenous or inherent uncertainty will manifest as looser confidence in the results.

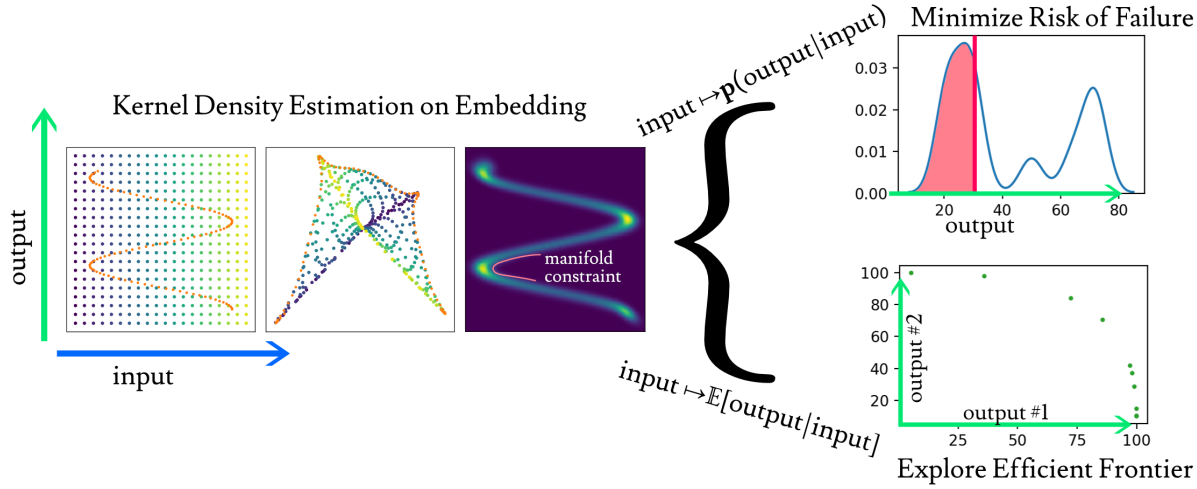
The effect of the inputs under consideration onto intrinsic build structure, namely porosity, is studied extensively on titanium alloys [11] with fruitful empirical results [13, 14]—note the avalanche of progress in 2020 on laser powder-bed fusion of different metals, including nickel alloy 718 [15], a variety of steels [16, 17, 18], and more titanium [19]. We investigate the additive manufacturing of AlSi10Mg, common in aerospace applications for which weight savings are of utmost importance [20]. Nevertheless, reproducibility needs dire improvement—especially in controlling the key physical metrics that this study scrutinized [21]. Porosity contributes significantly to strength, brittleness, and elongation, but so do a few other factors that have been studied sporadically. We make no distinction about the underlying causes for the observed material properties. Numerous of the aforementioned approaches [15, 11, 17] harness optical methods to monitor the evolution of the build microstructure and potentially optimize it online [22]. We focus on the context in which recording this data, or

the analysis thereof, is impractical or simply not implemented (as is the case in most of industry). We do not believe this is an artificial handicap to our study because the comprehensive and precise correspondence between micro-structural features (e.g. porosity) and macroscopic properties is still under active research [23, 22].

We present a novel framework that enables a completely data-driven quantification of empirical knowledge on the behavior of an AM system, and serves as a surrogate to expensive high-fidelity physical simulations. We strive to do so by approximating the full joint density of the build parameters (laser power, scan speed, and hatch distance) along with measured material properties that result from the fabrication process (ultimate tensile strength, yield strength, and elongation). The curse of dimensionality is overcome by first approximating the geometric manifold on which this set of variables lies.

All models are interpolation schemes for a finite empirical sample; choosing one over another reduces to trading inductive biases for those that suit a problem—described by its set of assumptions—best. In any physical system, only certain combinations of variables make sense. Usually the phenomena obey a specific form that may be captured by succinct fundamental principles, like a set of governing equations. Even when elegant laws are unattainable, we expect our observations to lie on an orderly subspace that may be characterized as a Riemmanian manifold immersed in the ambient Euclidean sample space [24, 25]. We chose to constrain our probability density estimate along this putative global structure. Through a change of variables we construct the one-way volume-preserving map to the embedding space. Superimposing Gaussian kernels onto each granular observation yields more accurate density estimates in the transformed space than in the original one. The basis for our model is kernel density estimation, which is nonparametric and thus canonical—unlike other statistical methodologies shy of artificial neural networks. Yet ours stands apart from the latter and from traditional nonparametric methods in its reduced sample-size requirement that is granted by our reliance on diffusion maps.

Our nonparametric formulation uses diffusion maps as a tool for delineating intrinsic geometric structure in the high-dimensional space, and liberates us from ill-justified stringencies that tend to come alongside other procedures in machine learning (such as restrictions to Gaussianity [4, 12] and polynomial fits [3]). To compute the density along a full grid in the sample space, we carefully violate an assumption in diffusion-maps literature in order to accommodate points that expressly do not lie on the manifold. The result consists of a powerful model that may be cross-sectioned in a plethora of ways to study complex interactions between parameters and outputs.



Overview of our methodology.

2 Method

Stochastic optimization is a formidable task, and its "holy grail" would be an accurate characterization of the output's probability density conditioned on any feasible input. A fundamental obstacle to explicit joint-density estimation between parameters and objectives, or inputs and outputs, is the curse of dimensionality. The intuitive notion behind Kernel Density Estimates (KDEs) falls apart as the sample-space dimension N rises in tune with model complexity. In this paper we will study a context in which (a) the underlying mechanism is unknown, calling for non-parametric models, and (b) evaluations of the

stochastic input-output relation are costly, so observations are few. To approach these restrictions in a high-dimensional formulation, we will grant one assumption that samples from the system lie on a Riemmanian manifold $\mathcal{M}$ of dimension $s \ll N$. We may then take advantage of this phenomenon by inducing a data-driven embedding of dimension $m \geq s$ with the samples on hand. Ozakin and Gray [26] provided theoretical results on the feasibility of a "submanifold" density estimate but garnering an approximation for the manifold itself was outside their scope. From the anthology of cutting edge nonlinear dimensionality-reduction techniques via manifold embeddings, including UMAP [27], we chose to employ the framework of diffusion maps [28, 29] for its physical interpretability and computational efficiency.

2.1 Overview of Diffusion Maps

To account for covariances and differences in scale between dimensions, since these are framed as realizations of variables with arbitrary characteristics, we employ the Mahalanobis distance $d : (\mathbb{R}^N \times \mathbb{R}^N) \to \mathbb{R}$:

$$d^2(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x} - \boldsymbol{y})^T C^{-1} (\boldsymbol{x} - \boldsymbol{y}) \tag{1}$$

between two points $\boldsymbol{x}, \boldsymbol{y}$ in Euclidean space according to some empirical covariance matrix C, elucidated below in Equation 3. Note that other metrics, even various norms L_p, could be used in place of the above definition. The diffusion-maps methodology proceeds with computing a geometric graph Laplacian stemming from a proximity (also known as "affinity" or "adjacency") matrix A among n points $X = [\boldsymbol{x}_1 \dots \boldsymbol{x}_n]$ composed of

$$a_{ij} = \exp\left(-\frac{d^2(\boldsymbol{x}_i, \boldsymbol{x}_j)}{\varepsilon}\right), \tag{2}$$

in the case of a Gaussian kernel with scaling factor ε. In diffusion maps, the Gaussian corresponds to the heat kernel and several elegant properties (which will be described below) follow from its use. It is also the most common kernel employed in density estimation, partly thanks to the distribution's emergence in nature from the central limit theorem, and also since it is the maximal entropy distribution given the first two moments (i.e. mean and variance). Now the covariance is estimated straightforwardly, furnished with the empirical means of each dimension $\boldsymbol{\mu} \in \mathbb{R}^N$, as

$$c_{ij} = \frac{1}{n} \sum_{k=1}^{n} (\boldsymbol{x}_k^{(i)} - \boldsymbol{\mu}^{(i)})(\boldsymbol{x}_k^{(j)} - \boldsymbol{\mu}^{(j)}) \tag{3}$$

considering each point in $\{\boldsymbol{x}_i\}_{i=1}^n$ to be an independent sample from the same distribution.

Coifman and Lafon [28] envisioned a family of diffusion maps parameterized by an $\alpha \in [0, 1]$ that controls density normalization as an exponent in the denominator. In the discrete scenario it is applied as follows:

$$w_{ij} = \frac{a_{ij}}{t_i^\alpha t_j^\alpha}, \text{ with } t_i = \sum_{k=1}^{n} a_{ik}.$$

The final step is to ascertain that W is row-stochastic by concocting a new matrix $P = R^{-1}W$ when R contains diagonal entries $r_{ii} = \sum_{k=1}^{n} w_{ik}$. Notice that P represents a Markov-chain transition matrix. It was shown [28] that $L = I - P$ (with I denoting the identity matrix of size n) converges to various kinds of relevant operators as $\varepsilon \to 0$, depending on the choice of α. The mechanics of α will be revisited later.

2.2 Extendable Manifold Coordinates

The singular vectors of P with the m largest eigenvalues*, which happen to correspond to those of L with the smallest singular values, provide the so-called diffusion maps coordinates $\Psi = [\lambda_1 \boldsymbol{\phi}_1 \dots]^T \in \mathbb{R}^{m \times N}$, with eigenvectors $\{\boldsymbol{\phi}_i\}$ and eigenvalues $\{\lambda_i\}$, of the samples in X. It is often convenient to extract the eigenvalues from the related $\widehat{P} = R^{-1/2} W R^{-1/2}$ that is symmetric and shares its eigenvectors with P. A plethora of literature examines the connection with Markov steady

*In practice, because the contribution to Euclidean distance scales with the eigenvalues, it is sufficient to select a large enough m such that the eigenvalues past that point have decayed to essentially zero. We guarantee this by keeping all positive eigenvalues.

states (i.e. equilibria) and the discovery that Euclidean distances in this new space in $\mathbb{R}^m$ (presumed to be a coordinate chart or embedding of $\mathcal{M}$) are tied to the geodesic distance on $\mathcal{M}$ according to all the possible random walks along the original samples.

Due to the loss of information through the mapping $\Phi : \mathbb{R}^n \to (\mathbb{R}^m \supseteq \mathcal{M})$, it is most likely impossible to obtain an inverse relation from the diffusion-maps coordinate space to the sample space. This shortcoming presents an obstacle in the context of kernel density estimation in $\mathbb{R}^n$ with respect to the computed embedding. In the present study we overcome this challenge by taking a grid in the sample space and supplying its points in bulk to the diffusion-maps algorithm (see Figure 1 for two toy examples), evaluating the kernel density estimate with respect to the points that putatively lie on the manifold at each of these now-warped grid points, and transferring the approximations of the joint-probability function back to their locations in the sample space. Hence the outlined scheme, streamlined yet nuanced, does not require a reverse mapping Φ^{-1}.

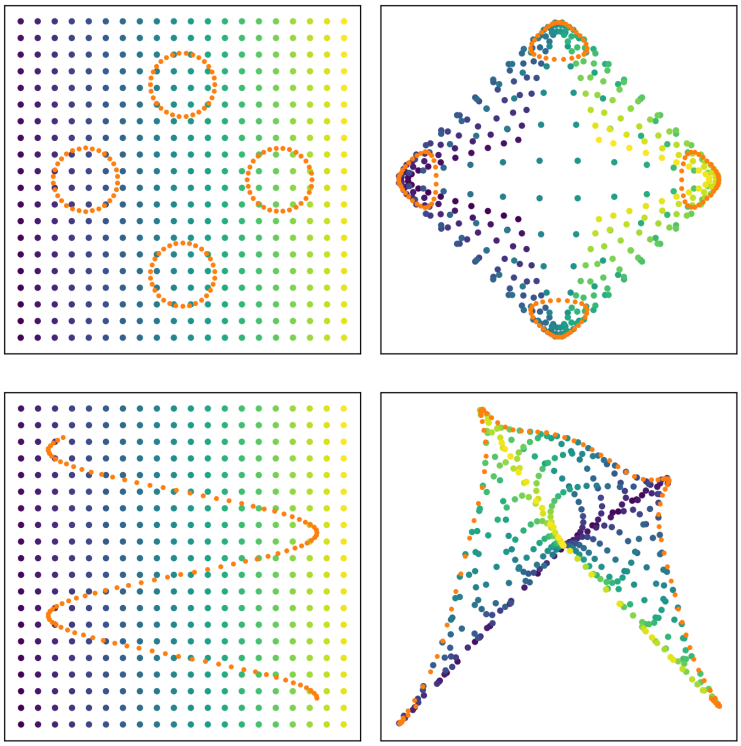


Figure 1: Illustrations of grid distortion via extended coordinates. Left: sample space. Right: derived embedding space. The set of orange dots in each row represent the two-dimensional points from which a manifold is approximated, and the multi-color grid is distorted to conform to that structure.

It is necessary to mutilate the above formulation to accommodate for two distinct sets of points: the reference set $\mathcal{R}$ and the auxiliary set $\mathcal{Z}$. $\mathcal{R}$ most closely resembles X in our original framing of diffusion maps, whereas $\mathcal{Z}$ will not be required to lie on the manifold; in fact, this is the first scrutiny of the technique in which the additional points explicitly do not come from $\mathcal{M}$ but rather an artificial grid.

A variety of applications have employed Coifman's formula to extend diffusion-maps coordinates [29, 30, 31, 32, 33] beyond the training set. What follows is an attempt to concisely explain it before delving into the modifications that were necessitated by our change of scope. Consider the same distance function as in Equation 1 but with refined notation to specify the source of the covariance matrix:

$$d^2_{\mathcal{R}}(\boldsymbol{r}, \boldsymbol{z}) = (\boldsymbol{r} - \boldsymbol{z})^T C^{-1}_{\mathcal{R}} (\boldsymbol{r} - \boldsymbol{z}).$$

Note that $\boldsymbol{r}$ and $\boldsymbol{z}$ may be drawn from either $\mathcal{R}$ or $\mathcal{Z}$. Define an anisotropic proximity $\widetilde{A}$ as

$$\tilde{a}_{ij} = \exp\left(-\frac{d^2_{\mathcal{R}}(\boldsymbol{r}_i, \boldsymbol{z}_j)}{\varepsilon}\right), \quad i = 1 \ldots |\mathcal{R}|, \; j = 1 \ldots |\mathcal{Z}|$$

where $\boldsymbol{r}_i \in \mathcal{R}, \boldsymbol{z}_j \in \mathcal{Z}$ are naturally indexed in their respective countable sets. It was shown [31] that $\widetilde{A}\widetilde{A}^T$ and A from Equation 2 (now referenced as $A_{\mathcal{R}}$) are *asymptotically* equivalent as $|\mathcal{R}| \to \infty$, assuming $\mathcal{Z}$ lies on the same manifold. Our scenario is no longer well-behaved because we lifted the imposition for the extension points to lie on $\mathcal{M}$, and as a consequence the perturbation-based proof of equivalence no longer holds. Thus our robust method is to calculate $A_{\mathcal{R}}$ and $\widetilde{A}$ separately. We scale $\widetilde{A}$ in a manner that stays coherent with the prior methodology:

$$\begin{aligned} \widetilde{W} &= \widetilde{A} \cdot \operatorname{diag}\left(\mathbf{1}^T \widetilde{A}\right)^{-\alpha}, \\ \widetilde{P} &= \widetilde{W} \cdot \operatorname{diag}\left(\mathbf{1}^T \widetilde{W}\right)^{-1/2} \end{aligned} \tag{4}$$

where through slight abuse of notation each of these lines normalizes the columns by their sum taken to a certain power. Given the reference diffusion eigenvectos $\Phi_{\mathcal{R}} = [\boldsymbol{\phi}_1 \dots \boldsymbol{\phi}_N] \in \mathbb{R}^{N \times N}$, we have the requisite tools to extend them to the auxiliary points $\widetilde{\Phi} = \widetilde{P} \cdot \Phi_{\mathcal{R}} \cdot \Lambda^{-1/2} \in \mathbb{R}^{|\mathcal{Z}| \times N}$. The new coordinates are given as $\widetilde{\Psi} = \Lambda \widetilde{\Phi}^T$, truncated to $\mathbb{R}^{m \times |\mathcal{Z}|}$.

Scales in the new space are slightly off from the original embedding, due to either numerical inaccuracies or approximation error (remains to be investigated). Our workaround is to augment $\mathcal{Z}$ with the reference points to garnish them with a modified set of coordinates as well, so that distances remain commensurable. This change entails the substitution of $\mathcal{Z}$ with the concatenated $\mathcal{Z}' = [\boldsymbol{r}_1 \dots \boldsymbol{r}_N \;\; \boldsymbol{z}_1 \dots \boldsymbol{z}_{|\mathcal{Z}|}]$ and the ensuing separation of $\widetilde{\Psi}$ into $\widetilde{\Psi}_{\mathcal{R}} = [\boldsymbol{\rho}_1 \dots \boldsymbol{\rho}_N]$ and $\widetilde{\Psi}_{\mathcal{Z}} = [\boldsymbol{\upsilon}_1 \dots \boldsymbol{\upsilon}_{|\mathcal{Z}|}]$.

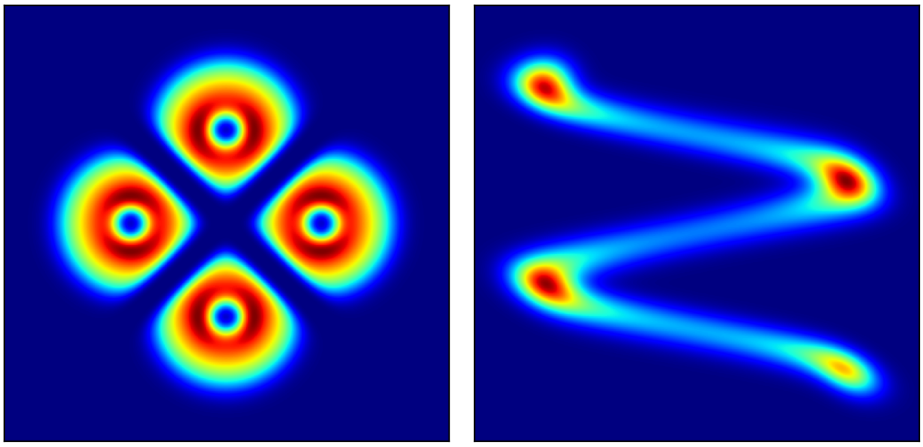


Figure 2: Illustrative kernel density estimates in line with the examples in Figure 1.

2.3 Kernel Density Scheme in Extended Diffusion Coordinates

Attempts heretofore on probabilistic reasoning along a manifold operated on a projection of the samples on the embedding coordinates [9]. An Îto stochastic differential equation represented a dissipative Hamiltonian system with "potential energy" governed by a Gaussian mixture on the original samples. By integrating this second-order differential equation within the projected embedding coordinates, the dynamical system was guided by the approximate manifold. Ghanem and Soize generated enough samples to perform a conditional Kernel Density Estimation (KDE) in the original Euclidean space [34, 35]. Now we sidestep the time-consuming simulation of a stochastic process by obtaining a KDE directly inside the embedding, thanks to our revised formula on the extension of coordinates. The dimensionality of the embedding space is potentially much lower; the lack of numerous artificial samples also relieves some of the computational burden of our stochastic optimization problem.

The most prevalent implementation of non-parametric multivariate density estimation makes use of the Gaussian kernel. The bandwidth term H, generally specified to scale the kernels in a manner akin to the Mahalanobis metric, admits the quadratic form $\boldsymbol{u}^T H^{-1} \boldsymbol{u}$. The determination of this matrix either involves a rule of thumb (like Silverman's or Scott's) or some flavor of cross validation. For the scope of testing our pioneering methodology we will commit to the simplification

of scaling uniformly: $H = hI$. Since density estimation occurs inside the embedding space, it would not be too egregious to make the Gaussian kernels perfectly symmetric along the diffusion-maps "dimensions". Ergo the estimation strategy is as follows:

$$f_{\widetilde{\Psi}}(\boldsymbol{v}_i; \widetilde{\Psi}_{\mathcal{R}}, \widetilde{\Psi}_{\mathcal{Z}}) \propto \sum_{j=1}^{N} \exp\left(\frac{\left\|\boldsymbol{v}_i - \boldsymbol{\rho}_j\right\|^2}{h^2}\right), \text{ where } \boldsymbol{v}_i \in \widetilde{\Psi}_{\mathcal{Z}}, \boldsymbol{\rho}_j \in \widetilde{\Psi}_{\mathcal{R}}. \tag{5}$$

Observe Figure 2 for an example of what this procedure entails, with a slightly exaggerated bandwidth to highlight the manifold dynamics. The two toy examples of Figure 1 are treated with the rest of the procedure due to Equations 5. One detail remains in order to complete the framework. Denote the shorthand form of the induced mapping $U_{\mathcal{R}} : \boldsymbol{z} \mapsto \boldsymbol{v}$ that is constructed from a pair $(\widetilde{\Psi}_{\mathcal{Z}}, \widetilde{\Psi}_{\mathcal{R}})$ that includes $\boldsymbol{z} \in \mathcal{Z}$. Then we are in search of the sample-space density $f(\boldsymbol{z})$ defined via the volume-preserving change of variables that follows naturally:

$$f_{\mathcal{Z}}(\boldsymbol{z}) = f_{\widetilde{\Psi}}\left(U(\boldsymbol{z})\right) \left|\det J_U(\boldsymbol{z})\right| \tag{6}$$

Where J_U is the Jacobian of U, which may also be denoted as $\frac{\partial \boldsymbol{v}}{\partial \boldsymbol{z}}$, evaluated at $\boldsymbol{z}$. Our intent, but not implementation, parallels that of [36]. Perrault-Joncas [37] showed that it is possible to directly compute the metric tensor $G_U = (J_U)^T J_U$ for tangent vectors, which is induced by the mapping U, when the Gaussian kernel is deployed. In brief, its pseudo-inverse—in our case we only care about the pseudo-determinant, which simplifies computation—can be related to certain manipulations on the Laplace-Beltrami operator. Recall that Equation 4 defines the extended family of operators $\widetilde{P} \in \mathbb{R}^{|\mathcal{Z}| \times N}$ parameterized by $\alpha \in [0, 1]$. Setting $\alpha = 1$ yields the Laplace-Beltrami approximation on the *reference points, evaluated at the auxiliary points* (notice the asymmetrical shape). We seek an operator defined wholly on the auxiliary points. The extended eigenvectors $\widetilde{\Phi} \in \mathbb{R}^{|\mathcal{Z}| \times N}$ are limited in number to the size of the reference set, yet by construction they serve as surrogate eigenvectors to the putative $\widehat{P} \in \mathbb{R}^{|\mathcal{Z}| \times |\mathcal{Z}|}$. We circumvented this bottleneck by defining $\widehat{P}$ on the auxiliary points by a low-rank approximation:

$$\widehat{P} = \widetilde{\Phi} \Lambda \widetilde{\Phi}^T$$

where Λ is obtained from Equation 4. The necessary Laplace-Beltrami operator is therefore available to calculate G_U via the procedure laid out in [37]. We lay it out briefly here: their fundamental result is that the quantity

$$\tilde{g}_{ij}(\boldsymbol{z}) = \frac{1}{2} \Delta_{\mathcal{M}} \left(U^{(i)} - U^{(i)}(\boldsymbol{z})\right) \left(U^{(j)} - U^{(j)}(\boldsymbol{z})\right)\Big|_{U^{(i)} = U^{(i)}(\boldsymbol{z}), U^{(j)} = U^{(j)}(\boldsymbol{z})}, \tag{7}$$

where superscript identifies scalar components of a vector and $\Delta_{\mathcal{M}}$ denotes the Laplace-Beltrami operator, forms a matrix of which the pseudo-inverse yields the metric G_U of the embedding with dimension s potentially greater than the intrinsic manifold dimension m. As a consequence, $\widetilde{G}$ may not be full-rank so we cannot take the direct inverse. We utilize the approximation of the aforementioned operator as by transforming Equation 7 into its discrete analog:

$$\tilde{g}_{ij}(\mathcal{Z}) = \frac{1}{2}\left(\widehat{L}U^{(i)}(\mathcal{Z})U^{(j)}(\mathcal{Z}) - U^{(i)}(\mathcal{Z})\widehat{L}U^{(j)}(\mathcal{Z}) - U^{(j)}(\mathcal{Z})\widehat{L}U^{(i)}(\mathcal{Z})\right)$$

where $\widehat{L} = \frac{4}{\varepsilon}(I - \widehat{P})$ and now $U^{(l)}(\mathcal{Z})$ terms are defined on the ordered set of points $\mathcal{Z}$ as

$$U^{(l)}(\mathcal{Z}) = [U^{(l)}(\boldsymbol{z}_1), \ldots U^{(l)}(\boldsymbol{z}_{|\mathcal{Z}|})]$$

and $\tilde{G}(\mathcal{Z})$ is an array of square matrices, or a 3-tensor with shape $m \times m \times |Z|$. The computational complexity of this subroutine is particularly expensive, and in our studies it appears that $|\det G_U| \approx 1$ and may not be practical or necessary to incorporate for an accurate KDE estimation.

With that in mind, the determinant of the Jacobian can be recovered as

$$|\det J_U| = \sqrt{|\det G_U|}.$$

We implicitly incorporate the formula in Equation 6. Suppose that $\mathcal{Z}$ is partitioned into two sets: $\mathcal{X} = \{z^{(1\ldots k)} | z \in \mathcal{Z}\}, \mathcal{Y} = \{z^{(k+1\ldots n)} | z \in \mathcal{Z}\}$. We remark that conditioning on a number of the latter variables in $\boldsymbol{z}$ may proceed as

$$f_{\mathcal{X}|\mathcal{Y}}(z^{(1)}, \ldots z^{(k)} | z^{(k+1)}, \ldots z^{(n)}) = \frac{f_{\mathcal{Z}}(z^{(1)}, \ldots z^{(k)},\ z^{(k+1)}, \ldots z^{(n)})}{\sum_{q_1} \cdots \sum_{q_k} f_{\mathcal{Z}}(q_1, \ldots q_k,\ z^{(k+1)}, \ldots z^{(n)})}, \tag{8}$$

where the constant multiple depends on the discretization scheme used in the denominator to evaluate the marginals. We also note that the order of variables is arbitrary and we chose to condition on the latter portion for no particular reason. This formulation is especially valuable because we aim to employ the framework in stochastic optimization by obtaining a diffusion-maps kernel density estimate of the joint distribution of a system's inputs and outputs, then either condition on specific outputs to predict the requisite inputs (and the feasibility thereof) or more classically condition on inputs to quantify our knowledge of the possible outputs. The concrete optimization strategy will be elaborated further on. We wish to base our critique of the model's accuracy on the conditional expectations it produces. Therefore in probabilistic terms, the variable encapsulating the observable parts of the system's behavior is termed $\boldsymbol{S} : \Omega \to \mathbb{R}^n$ for an unknown measurable set Ω and induces the mapping $\omega \mapsto \boldsymbol{X}_1(\omega) \dots \boldsymbol{X}_k(\omega), \boldsymbol{Y}_{k+1}(\omega) \dots \boldsymbol{Y}_n(\omega)$. This notation allows us to write the conditional expectation of the inputs as an adaptation of Equation 8:

$$\begin{aligned} \mathbb{E}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}_1 \dots \boldsymbol{X}_k | y^{(k+1)} \dots y^{(n)}] &= \sum_{q_1} \cdots \sum_{q_k} [q_1, \dots q_k] f_{\boldsymbol{X}|\boldsymbol{Y}}(q_1, \dots q_k | y^{(k+1)}, \dots y^{(n)}) \\ &= \frac{\sum_{q_1} \cdots \sum_{q_k} [q_1, \dots q_k] f_{\boldsymbol{X},\boldsymbol{Y}}(q_1, \dots q_k,\ y^{(k+1)}, \dots y^{(n)})}{\sum_{q_1} \cdots \sum_{q_k} f_{\boldsymbol{X},\boldsymbol{Y}}(q_1, \dots q_k,\ y^{(k+1)}, \dots y^{(n)})}, \end{aligned} \tag{9}$$

and analogously for the conditional expectation of the outputs. The reference set $\mathcal{R}$ consists of the available realizations $\boldsymbol{r}_i = \boldsymbol{S}(\omega_i)$ for $i \in 1 \dots N$. Note that Equation 9 relies not on the constant factors that were introduced implicitly in Equation 5 because it is normalized by the denominator.

2.4 Cross-validating Conditional Expectation

Let $\mathcal{R}^{-l}$ denote $\mathcal{R} \setminus \{\boldsymbol{r}_l\}$ and $\mathbb{E}^{-l}$ the expectation with a substitution of $\mathcal{R} \mapsto \mathcal{R}^{-l}$ in Equation 5. Then

$$\boldsymbol{r}_l + \boldsymbol{\epsilon}_l = \left(\mathbb{E}^{-l}_{\boldsymbol{X}|\boldsymbol{Y}}[\boldsymbol{X}_1, \dots \boldsymbol{X}_k | y_l^{(k+1)}, \dots y_l^{(n)}], \mathbb{E}^{-l}_{\boldsymbol{Y}|\boldsymbol{X}}[\boldsymbol{Y}_{k+1}, \dots \boldsymbol{Y}_n | x_l^{(1)}, \dots x_l^{(k)}] \right) \tag{10}$$

provides an avenue for k-fold cross validation, where $k = N$. In this case we seek to find a bandwidth h as well as crucial embedding parameters $\boldsymbol{\theta} = (\varepsilon, \alpha)$ that minimize the overall normalized square error of the cross-validated conditional expectation:

$$\hat{h}, \hat{\boldsymbol{\theta}} = \arg\min_{h,\boldsymbol{\theta}} \sum_{l=1}^{n} \sum_{i=1}^{N} \frac{\epsilon_{l,i}^2}{\sigma_i^2}.$$

Here σ_i^2 denotes the empirical variance of the variable corresponding to the ith dimension, also written as $\mathrm{Var}(\boldsymbol{S}_i)$. Our reasoning follows the standard method of bandwidth selection in kernel regression as detailed in [38]. In the following sections we decompose $\boldsymbol{\epsilon}$ in order to consider the normalized error with respect to solely the conditional inputs, and then to the conditional outputs, and contemplate the cause and effect of any observable differences. We term the forward cross-validation the metric on the prediction error of the outputs, and likewise the backward cross-validation for predicted inputs.

In fact, if we are to assume that the outputs are mapped from the inputs via some measurable function $\boldsymbol{g}$, they may be written like so:

$$\boldsymbol{g}(\boldsymbol{X}_1, \dots \boldsymbol{X}_k) = (\boldsymbol{Y}_{k+1}, \dots \boldsymbol{Y}_n)\,.$$

Now the system state $\boldsymbol{S}$ contains k independent variables and $n-k$ dependent ones. The image of $\boldsymbol{g}$ is generally unknown, i.e. we cannot choose arbitrary outputs and expect the estimated inputs to bear any meaning. We may explore it by feeding in various input realizations and observing the outputs—in our procedure this would manifest as conditioning on the inputs. Alternatively, we could estimate the marginal cumulative distribution of the outputs and trace the boundaries of the most optimal configurations that do not drop below a certain confidence threshold; we would thus arrive at an efficient (Pareto) frontier, albeit stochastic. In the present study, we resorted on solely conditioning on the inputs.

3 Results from a Case Study

The economics of air and space vehicle weight reduction provides a compelling reason to seek new methods of design and manufacture for structural component mass optimization. NASA estimates that it costs over $10,000 to put a pound of

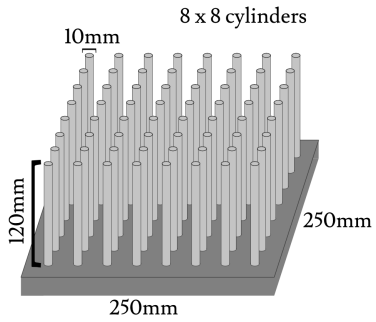


Figure 3: The build plate layout. The actual build file was deemed proprietary, so this illustration is provided instead. Dimensions are to scale.

payload in Earth orbit [39]. For commercial aircraft, a Boeing report from 2012 [40] suggests that removing just one pound of weight from each aircraft in a fleet of 619 commercial aircraft could save about 11,000 gallons of fuel annually. Laser powder-bed fusion, a layer-wise fabrication process in which structural components are fabricated layer by layer with processed digital information, provides the opportunity to lightweight structural components in ways that are not possible with traditional manufacturing methods.

In addition to manufacturing components that have been optimized for minimum mass, AM also provides the ability to increase the functionality of the component (e.g., incorporate internal cooling channels), increase its complexity, consolidate a system of parts into a single component for volumetric efficiency and reduced manufacturing, assembly, warehousing, and serializing, and significantly reduce manufacturing waste. With respect to the latter, additive manufacturing is associated with 10–20% waste material compared to 80-90% of the original billet for conventional processes, which is significant, especially considering the high-priced materials used in the aerospace industry such as titanium alloys.

Furthermore, AM exploits the solutions of topology optimization methodologies, providing lightweight structural components that cannot be, or at best may be difficult to be, manufactured traditionally. Despite the many advantages of AM afforded to the aerospace industry, there still exists a palpable apprehension and conservatism among many industry leaders regarding the readiness of the manufacturing modality for flight parts. Because of the high stakes involved (e.g., safety of human life, and cost of spacecraft or aircraft), the aerospace industry requires assurances that parts designed for and manufactured by AM processes are reliable and structurally sound. Hence the ability to predict the process parameters that result on optimum mechanical properties is of chief interest.

3.1 Experimental Setup

All components in this study were fabricated from aluminum alloy AlSi10Mg powder in an EOS M290 machine[1]. Fifty-one unique parameter sets consisting of combinations of laser power, scan speed, and hatch distance were created and assigned to tangible tensile test coupons, which were sent to an outside laboratory to measure ultimate (tensile) strength, yield strength and elongation. The coupons were fabricated as vertical cylindrical bars, and subsequently machined into the dog-bone shape according to the ASTM E8 standard. Each coupon was the same dimension (see Figure 3) and fabricated on the same machine in an effort to eliminate machine-to-machine and geometry-to-geometry variability. The 51 input-output pairs under scrutiny are similar to those that previously paved the way towards real aerospace-grade parts.

The coupons were printed with the intent to be destructively tested, serving as an inexpensive proxy for the validation of actual build components. The particles comprising the metal powder were characterized by percentiles on their diameters; the median (50th percentile) was 42.78 microns and the 90th percentile was 69.88 microns. The layer thickness throughout was 60 microns and the laser focus diameter 100 microns, while the hatch spacing (simply termed hatch) along with the laser power (termed power) and scanning speed (termed speed) varied according to an experimental design driven by an expert's domain knowledge. All values are normalized between typical, yet arbitrary, minima and maxima of each variable for the purpose of protecting sensitive information. Formally the specific combination of chosen parameter values was arbitrary; thus an effective inference method would require robustness to biased sampling. Other factors that influence the build

process and could not be held constant were considered unobservable. Their effect is manifested implicitly as statistical scatter in our inference. The system to be modeled was hence a mapping from the three parameters to the three test results, incorporating noise in the form of measurement error and variation in build artifacts intrinsic to the powder-bed process.

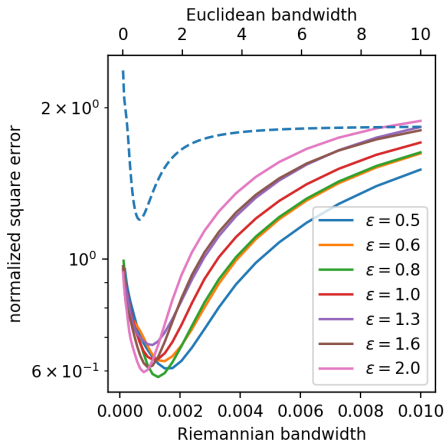


Figure 4: Cross-validation grid search with comparison to Euclidean KDE. The cluster of curves corresponds to errors of our Riemannian (diffusion-maps) KDE procedure with varying kernel scales ε (α is fixed to 1) and the clearly separated dashed curve presents the results of our control, the Euclidean KDE. The green curve appears to come from the best fit. Normalized square errors are plotted on a logarithmic scale.

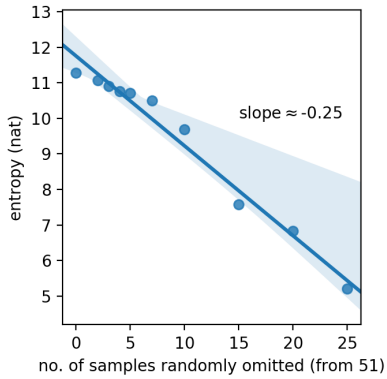


Figure 5: Entropies of the joint density after leaving out various amounts of samples, shuffled each time. Linear fit with 99%-confidence intervals is superimposed. It appears that each sample contributes a roughly constant 0.25 nats to the overall entropy of the model.

3.2 Analysis: Inference and Optimization

We then performed the aforementioned cross-validation technique (see Equation 10) to tune the model parameters, which were explicitly the diffusion-maps kernel size ε and the kernel-density bandwidth h. We maintained $\alpha = 1$, enforcing that we approximated the Laplace-Beltrami operator along the manifold. This operator is, asymptotically, agnostic to density (i.e. the sampling distribution on which we place no hard assumptions in this setting and which is also of no interest to us) and should gauge solely the geometric structure of the observations. In the field of machine learning, there is a strict distinction between the training set, validation set, and testing set. A typical machine-learning model is composed of numerous parameters, each of which are estimated with the help of a combination of hyper-parameters, e.g. the learning rate plugged into the gradient-descent algorithm. The aptly-named training set is employed for the learning algorithm to train the model; "peeking" into the validation set is allowed for the sake of hyper-parameter selection. Finally, since the validation set was

also used to shape the resultant model, a testing set exists for the ultimate evaluation. Following this terminology, k-fold cross validation is a resourceful training/validation set mixture where all the samples end up playing a part in training and validation. Every time a validation sample is selected, a diffusion-maps embedding is learned and evaluated. In this case the kernel size and KDE bandwidth are considered hyper-parameters that the validation scheme aids in selecting. An analog to a testing set is the predicted optimal values for build parameters, that we then tested physically.

We cross-validated both directions of the system: the expectation of the build parameters (inputs) conditioned on the mechanical properties (outputs) and vice versa. In computing the mean normalized square error, we took note of the apparent phenomenon that it is more difficult to predict outputs based on inputs than to predict inputs based on outputs. This is likely due to the higher entropy present in the measured quantities that are subject to multiple sources of noise, yet it goes contrary to the intuition that the input-output mapping is not necessarily invertible; i.e., some combinations of outputs may be produced from different input configurations but not vice versa (discounting scatter). Figure 4 shows the results of grid search on the forward (i.e. input-to-output) cross validation with a logarithmic discretization of the bandwidths h and varying kernel scales ε. Here, the error curves of the Euclidean KDE versus the diffusion-maps KDEs are clearly separated; their best-case normalized square errors are about 1.2 and 0.6, respectively. Cross validation in the backwards direction does not exhibit such a stark separation—both errors in that instance are around 0.5. We care about prediction in the forward direction, so upon that basis we selected the hyper-parameters $\varepsilon = 0.8$ and $h = 0.00123$. The bandwidths between Euclidean and diffusion-maps (i.e. Riemannian) kernels are incommensurable, but we noticed visually that the cross-validated densities in the latter are sharper and therefore of lower entropy.

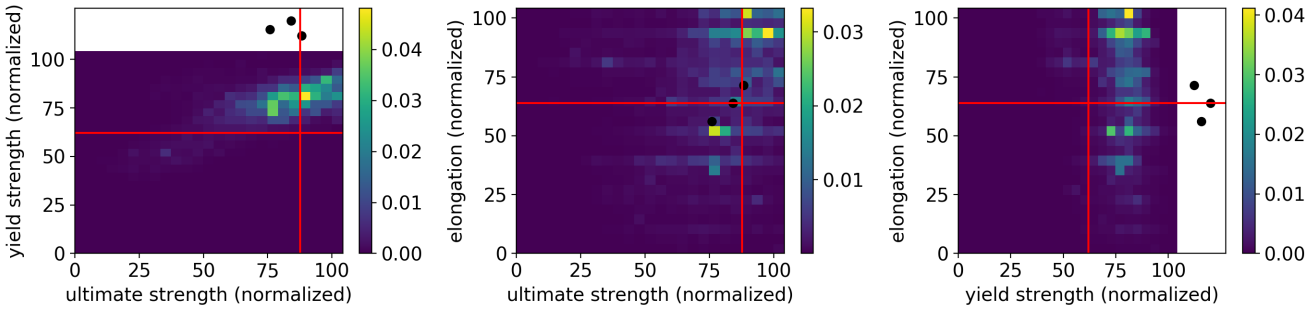
We sought an understanding of the effect of sample size by presenting Figure 5—short of launching separate full-fledged experiments for different levels of subsampled data. It enables the argument that the amount of information produced by each additional sample (and hence speculative improvement to the quality of the prediction) is constant, though it slightly tapers off after the introduction of roughly the 45^{th} sample. Specifically we tracked entropies with progressive subsampling (akin to bootstrapping) and reevaluated the full joint diffusion-maps kernel density quantized on a grid of side-length twelve. Figure 4 and the subsequent searches in Sections 3.2.1 & 3.2.2 operated on the conditional density function, which accommodates much larger grids, although they were kept at length ten for the sake of computing as many random samples as possible. Again, the top of the curve hints of a budding inflection point—a potential sign of increasing informational redundancy.

For the sake of stochastic optimization, a few avenues were open for exploration. We could attempt to trace the image of the system via contours on the cumulative (marginal) density of the outputs, to garner confidence bounds for feasible outputs and then conditioning on desirable combinations to gauge the uncertainty of the hypothetical inputs to produce them. That methodology is cumbersome and is therefore left as the subject of future investigations. Instead, we employed a Monte Carlo search on the set of possible inputs, which is known a priori. In our basic case it was trivial—a bounding box with a small margin around the extrema of the dataset. More precisely, the allowed range in each dimension (both for the estimation of density and permitted values in the optimization scheme) was 0.1 empirical standard deviations from the observed minimum and maximum.

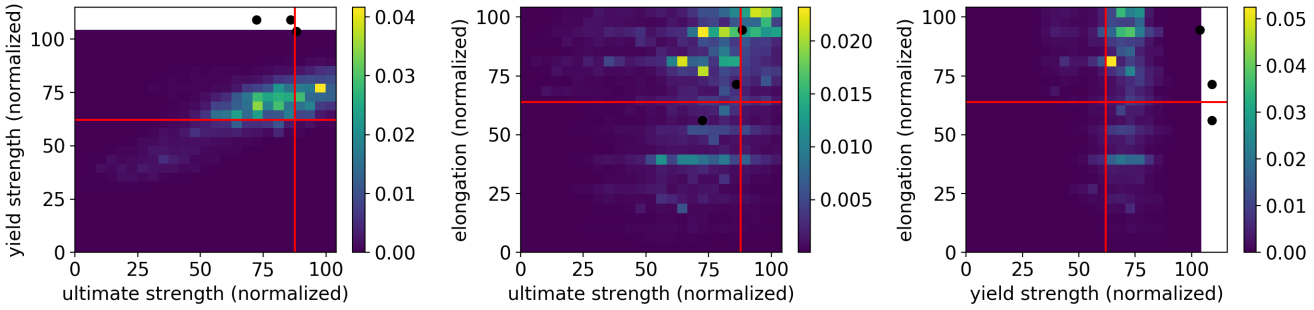
Domain experts in additive manufacturing seek to understand the full effects of the build process on the structural integrity of a desired product. Our proposed methodology for joint density estimation within a diffusion-maps embedding lends itself well to such a task because it approximates the full probabilistic character of the quantities of interest. In a safety-critical context like those within the aerospace and defense industries, it is crucial to identify the configuration of parameters that minimizes the chance of failure or equivalently maximizes the chance of an acceptable outcome. In the following section we showcase the results of such an investigation and note that the resulting build parameters are a prediction because they do not exist in the data. According to the inferred model, they are safer than any of the builds tested heretofore.

3.2.1 Maximizing safety

Figures 6a and 6b exhibit the result of our procedure using insider-defined thresholds for (normalized) ultimate strength, yield strength, and elongation of 87.7%, 62.0%, and 63.8% respectively. Notice that since we employ a Monte Carlo optimization scheme the solutions are slightly different, though surprisingly stable. Yield is most strongly held within the defined boundaries. We attempted to relax the constraint on ultimate strength from 87.7% to 75.9% in an attempt to give more leeway on the objective function. Surprisingly, the resultant build parameters are nearly identical to those of the other two trials.



(a) Normalized hatch of 7.37%, speed of 46.4%, and power of 2.08%.



(b) Normalized hatch of 14.8%, speed of 36.5%, and power of 6.64%.

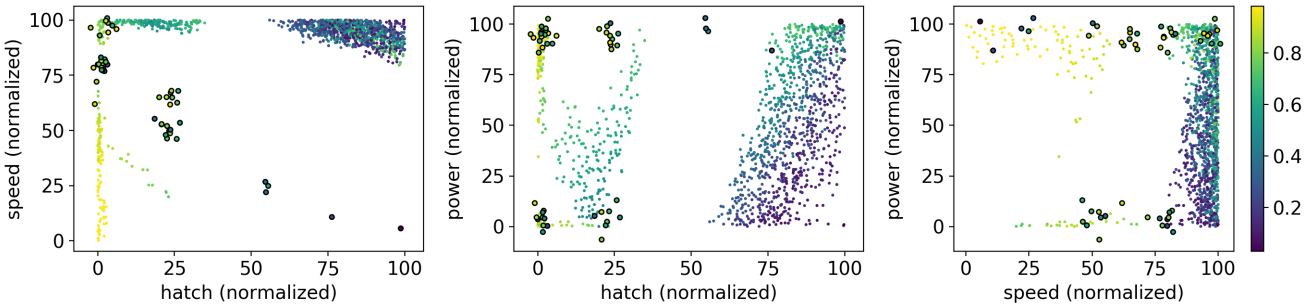
Figure 6: Each figure displays the result of a search through 100,000 random input combinations to find the safest configuration given certain output thresholds. Black dots present our attempt at out-of-sample validation that is discussed in Section 4.1.

3.2.2 Finding an efficient frontier of expectations

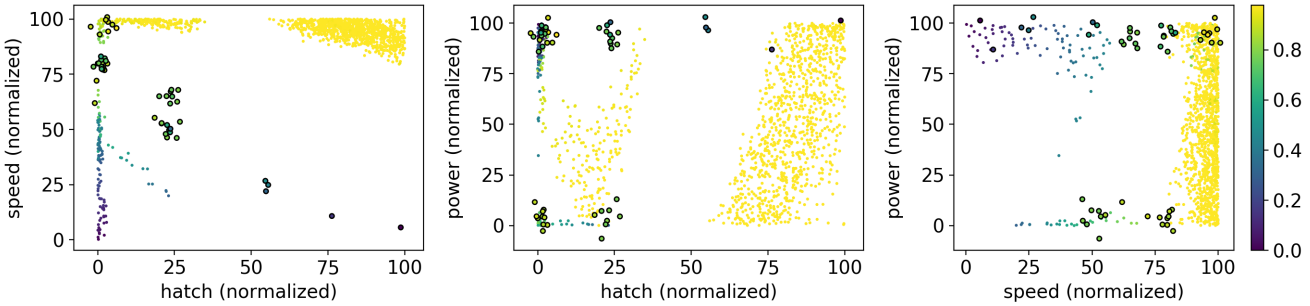
In our wish to attain a complete perspective on the dynamics that drive the system under study, it is valuable to illuminate the tradeoffs inherent in the design of the structure. We switched from the study based on quantiles of the "safe" region, as dictated by domain experts, to a multi-objective optimization scheme on conditional expectations of the material properties. We lose information on certainty (the greatest advantage offered by our methodology) along the way. In exchange it is easier to explore the Pareto frontier: the set of input-output pairs that are efficient with respect to the estimated outputs. It is important to keep in mind that we are not looking at the tails of the output distribution, but at the extrema of the expected values. To aide in the visualization of this high-dimensional space we display all the noteworthy cross sections along the inputs, with coloring according to the expected outputs, in Figure 7 and along the outputs in Figure 8. 50,000 trials resulted in 1,289 identified points on the frontier. Both sets of charts include as reference the data itself. In Figures 7 the inputs of the data are perturbed by a small amount of Gaussian noise for the purpose of visualization, since otherwise too many points would overlap on highly repeated values.

4 Discussion

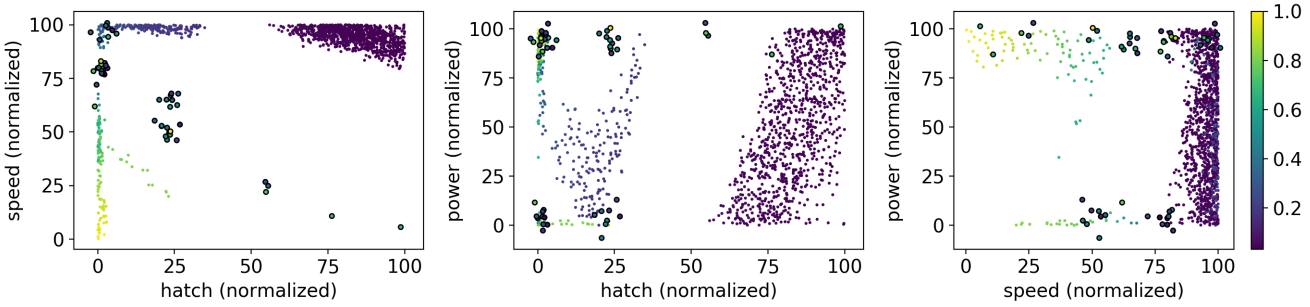
We demonstrated empirically in Section 3.2 the vast increase in accuracy that is induced by the transformation of the ambient sample space to the diffusion-maps embedding guided by a limited amount of system realizations. Furthermore, in Section 3.2.1 we obtained a close-to-optimal set of input build parameters that minimizes the chance of a failure event, herein defined as a manufactured product with one of its tested material properties falling below a specific threshold. These guidelines were crafted by current practices in the additive-manufacturing industry with metals. The estimated safety margins for yield strength surpassed expectations, as they did for elongation to a lesser extent. It appears that the model may exaggerate the possibility of failure by means of subpar ultimate strength because, given that it only saw the small number of samples we gave it, it "had to" err on the side of uncertainty. The predicted probability of an acceptable ultimate strength lies in



(a) Ultimate strength (normalized).



(b) Yield strength (normalized).



(c) Elongation (normalized).

Figure 7: Normalized inputs associated with the simulated frontier. Experimental data are marked with a black circle outline.

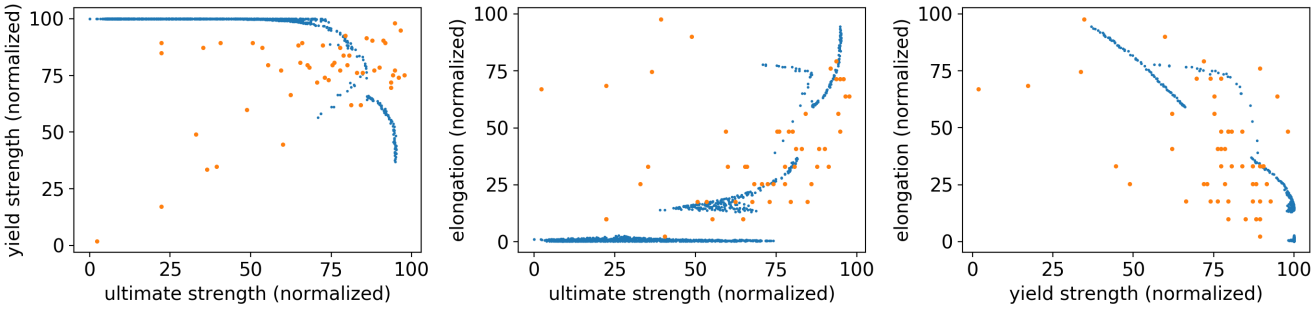


Figure 8: Outputs (normalized). Blue is the inferred (simulated) frontier and orange is the experimental data.

the range of 65-70%, which is significantly lower than what has been achieved empirically. The densities themselves were constructed on the basis of accuracy in predicting the existing data (via cross validation), so perhaps the procedure found it necessary to increase entropy to reach and fill in empty regions of the $(n-1)$-sample space. For this reason, we would make

the case that this mismatch of predicted uncertainty with what has been seen in practice would not be grounds for dismissing the parameters that the model claims are safest. We also remark that since it is optimizing in the 3-dimensional space of ultimate strength, yield strength, and elongation, it found it worthwhile to sacrifice the certainty on ultimate strength for a yield strength that is 97% of the time higher than the acceptable bounds and an elongation that, given its unwieldy physical nature, is adequately controlled (though with a marginal chance of success in a similar range as that of ultimate strength, specifically 65-70%).

The results of our analysis insinuated that the solution set from the optimization problem presented in Figure 6 minimized the chance of producing a failed tensile coupon, with high probability, *compared to any other choice of parameters*. We explored the validity of this assertion in Section 4.1. Typically, in the aerospace setting, a decently higher hatch size and laser power are employed: namely a hatch size of 23%, laser speed of 68%, and laser power of 96% (which is bimodal, so either close to 100% or to 0%) in contrast to the settings in Figure 6. Eight controls were tested with this typical setting; the mean elongation of the parts from Figure 6a was 1.27 higher than that of the control group, and the standard deviation was about the same. Similarly, Figure 6b reported an improvement in elongation by a factor of 1.39. As for ultimate tensile strength, the means were statistically equivalent across the board but Figure 6a achieved a standard deviation 0.50 the size of that for the controls. Likewise, Figure 6b experienced a reduction to about 0.68 that of the controls.

We would like to note that the present framework allows for the optimization of any function of the outputs. In the instance of minimizing the chance of failure, the definition of an acceptable output combination is not restricted to a bounding box as was utilized in the current study. The hard, independent thresholds were a product of current practice in the aerospace industry. Perhaps a more nuanced heuristic could be developed in the future—one that penalizes for the total time necessary to print (i.e. the reciprocal of hatch size times speed), for instance. This heuristic would answer the question posed above of whether

The model does not require any iterative training, and is thus rather efficient. Searching the resultant conditional density space takes up the bulk of computational resources and its demands scale directly with the dimensionality of the parameter space. As noted above, the computation of the metric tensor in Equation 7 is the most costly part of the procedure and could perhaps be approximated more efficiently, depending on the criticality of that component.

The second portion of our scrutiny—Section 3.2.2—involved the identification of an efficient frontier on the expectations of the three outputs. Figure 8 thereof lends itself well to an interpretation that aligns with first principles. Following the rough perceivable trends to their theoretical conclusions, it appears that ultimate strength and yield strength are largely independent of each other (indicated by the frontier that traces an angle close to 90°). Then with a sufficiently high ultimate strength, it begins to correlate positively with elongation. Circularly, a sufficiently high elongation relates negatively to yield strength. One could surmise that a high yield strength induces brittleness in the material that deteriorates its flexibility. An ultimate strength much higher than yield strength, suggesting the material can undergo a large amount of strain in which it exhibits plasticity, seamlessly fits the narrative based on varying degrees of brittleness. We are delighted that the data-driven formulation appears—upon elementary inspection—to suit the underlying physics that govern the system.

4.1 Out-of-sample validation

Coupons furnished with hatch, speed, and power corresponding to the putative optimals of 6a and 6b were printed in an identical setting to that which produced the original data set, and were tested at the same facility hired previously. The three measurement results from each set of build parameters are plotted in the above figures. Ultimate strength and elongation roughly follow the prescribed distributions. Yield strength is idiosyncratic. For the sake of easy diagnosis we had printed a "control" coupon with typical build parameters and also a few selected from Figures 7 that were supposed to yield abysmal characteristics. The control, of which the parameters were tested numerous times in the original dataset, also produced an abnormally high yield. The purposely negative coupons did not. Since measured yield strength per se was idiosyncratic in both the control and the optimal set of coupons, we attribute this inconsistency to some weakness in the protocol engaged by the testing facility. Yield is indeed the most difficult of the three measurements to quantify from empirical stress-strain curves.

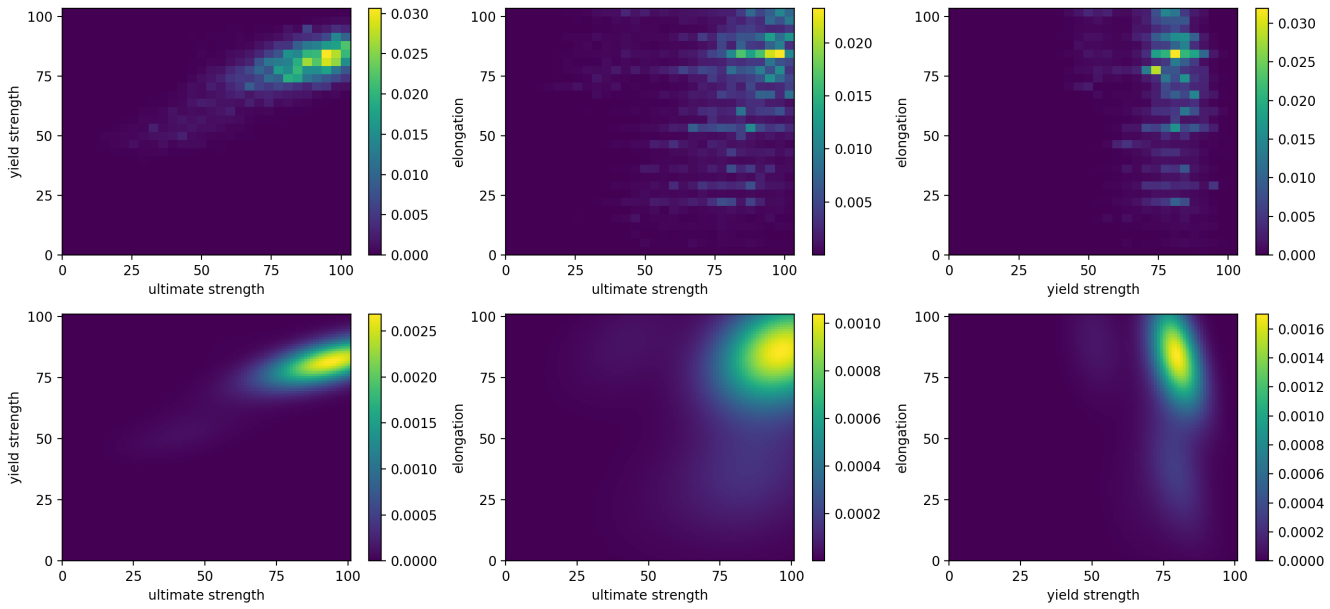


Figure 9: Optimization conducted on smooth densities, with resultant parameters akin to those of Figure 6. Top: Jacobians included a posteriori. Bottom: Jacobians omitted.

4.2 Future Concerns & Further Study

Throughout the investigation we were hindered by the computational complexity of approximating the diffusion-maps Jacobians in order to preserve volume across the mapping. We also realized that they were the source of apparent roughness/jaggedness of the estimated conditional probabilities. This shortcoming was especially egregious when time constraints necessitated that we keep the grid length at each dimension to 10 discrete points during the Monte Carlo searches. As shown in Figure 9, discarding scaling effects of the Jacobian's determinant introduced vastly smoother densities. The well-behaved surfaces made it safer to discretize at low resolution, and they also increased efficiency. We note anecdotally that the Kullback-Leibler divergences between smooth and "abrasive" approximations on low-resolution grids tend to be small. Luckily, the instability of the approximate Jacobians did not impede on the results' viability in this particular study.

One reviewer requested a display of some sort of sensitivity of the conditional output distributions in Figure 6 to the particular choice of input. We are intrigued by this line of thought: could we derive a first-order Taylor approximation of the gradient of a divergence measure with respect to the input parameters? That would allow us to conduct a proper sensitivity analysis on our solution to the Monte Carlo optimization. Evidently, a rigorous treatment of the propagation through the diffusion-maps embedding is necessary.

5 Conclusion

The presented technique is made possible by a stochastic interpretation combined with a few clever tricks applied to the framework of diffusion maps. The resultant formulation produced a flexible model of the interplay between build parameters in the laser powder-bed fusion process for additive manufacturing with metals. Whereas further validation and refinement is necessary before this novel methodology can make an appearance as a practical tool in industry, its power has been revealed.

6 Final Remarks

6.1 Disclosure for Possible Conflicts of Interest

Morf3D, Inc financed the empirical case study. Development and implementation of the algorithm was performed independently by the two authors.

6.2 Data Availability

The data sets analyzed during the current study are not publicly available due to intellectual-property concerns but normalized versions thereof are available from the corresponding author upon reasonable request.

6.3 Acknowledgements

The authors would like to acknowledge the invaluable advice and generous support of Melissa E. Orme, who was the Chief Technology Officer at Morf3D, Inc when the present study took place.

6.4 Licensing

References

[1] L. Bian, N. Shamsaei, and J. M. Usher, *Laser-based additive manufacturing of metal parts.* CRC Press, 2018.

[2] J. Beuth and N. Klingbeil, "The role of process variables in laser-based direct metal solid freeform fabrication," *Journal of The Minerals, Metals & Materials Society*, vol. 53, no. 9, pp. 36–39, 2001.

[3] K. Aoyagi, H. Wang, H. Sudo, and A. Chiba, "Simple method to construct process maps for additive manufacturing using a support vector machine," *Additive Manufacturing*, vol. 27, pp. 353–362, 2019.

[4] I. A. Okaro, S. Jayasinghe, C. Sutcliffe, K. Black, P. Paoletti, and P. L. Green, "Automatic fault detection for laser powder-bed fusion using semi-supervised machine learning," *Additive Manufacturing*, vol. 27, pp. 42–53, 2019.

[5] J. Ling, M. Hutchinson, E. Antono, S. Paradiso, and B. Meredig, "High-dimensional materials and process optimization using data-driven experimental design with well-calibrated uncertainty estimates," *Integr Mater Manuf Innov*, vol. 6, pp. 207–217, 2017.

[6] A. Hunter, B. A. Moore, M. Mudunuru, V. Chau, R. Tchoua, C. Nyshadham, S. Karra, D. O'Malley, E. Rougier, H. Viswanathan, and G. Srinivasan, "Reduced-order modeling through machine learning and graph-theoretic approaches for brittle fracture applications," *Computational Materials Science*, vol. 157, pp. 87–98, 2019.

[7] B. A. Moore, E. Rougier, D. O'Malley, G. Srinivasan, A. Hunter, and H. Viswanathan, "Predictive modeling of dynamic fracture growth in brittle materials with machine learning," *Computational Materials Science*, vol. 148, pp. 46–53, 2018.

[8] M. Schwarzer, B. Rogan, Y. Ruan, Z. Song, D. Y. Lee, A. G. Percus, V. T. Chau, B. A. Moore, E. Rougier, H. S. Viswanathan, and G. Srinivasan, "Learning to fail: Predicting fracture evolution in brittle material models using recurrent graph convolutional neural networks," *Computational Materials Science*, vol. 162, pp. 322–332, 2019.

[9] C. Soize and R. Ghanem, "Data-driven probability concentration and sampling on manifold," *J. Comput. Phys.*, vol. 321, pp. 242–258, 2016.

[10] A. Rovinelli, M. D. Sangid, H. Proudhon, and W. Ludwig, "Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials," *npj Computational Materials*, vol. 4, no. 35, 2018.

[11] F. Imani, A. Gaikwad, M. Montazeri, P. Rao, H. Yang, and E. Reutzel, "Process mapping and in-process monitoring of porosity in laser powder bed fusion using layerwise optical imaging," *Journal of Manufacturing Science and Engineering*, vol. 140, no. 10, 2018.

[12] A. Olleak and Z. Xi, "Calibration and validation framework for selective laser melting process based on multi-fidelity models and limited experiment data," *Journal of Mechanical Design*, vol. 142, no. 8, 2020.

[13] H. Gong, K. Rafi, H. Gu, T. Starr, and B. Stucker, "Analysis of defect generation in ti-6al-4v parts made using powder bed fusion additive manufacturing processes," *Additive Manufacturing*, vol. 1-4, pp. 87–98, 2014.

[14] C. Qiu, G. A. Ravi, and M. M. Attallah, "Microstructural control during direct laser deposition of a β-titanium alloy," *Materials and Design*, vol. 81, pp. 21–30, 2015.

[15] M. Montazeri, A. R. Nassar, A. J. Dunbar, and P. Rao, "In-process monitoring of porosity in additive manufacturing using optical emission spectroscopy," *IISE Transactions*, vol. 52, no. 5, pp. 500–515, 2020.

[16] Y. Zhou, S. Chen, X. Chen, J. Liang, C. Liu, and M. Wang, "The effect of laser scanning speed on microstructural evolution during direct laser deposition 12crni2 alloy steel," *Optics and Laser Technology*, vol. 125, p. 106041, 2020.

[17] J. A. Mitchell, T. A. Ivanoff, D. Dagel, J. D. Madison, and B. Jared, "Linking pyrometry to porosity in additively manufactured metals," *Additive Manufacturing*, vol. 31, p. 100946, 2020.

[18] S. Shamsdini, S. Shakerin, A. Hadadzadeh, B. S. Amirkhiz, and M. Mohammadi, "A trade-off between powder layer thickness and mechanical properties in additively manufactured maraging steels," *Materials Science & Engineering A*, vol. 776, p. 139041, 2020.

[19] X. Zhou, N. Dai, M. Chu, L. Wang, D. Li, L. Zhou, and X. Cheng, "X-ray ct analysis of the influence of process on defect in ti-6al-4v parts produced with selective laser melting technology," *Int J Adv Manuf Technol*, vol. 106, pp. 3–14, 2020.

[20] M. E. Orme, R. Vernon, I. J. Madera, R. Yancey, and F. Mouriaux, "Additive manufacturing of lightweight, optimized, metallic components suitable for space flight," *Journal of Spacecraft and Rockets*, vol. 54, no. 5, pp. 1050–1059, 2017.

[21] L. Dowling, J. Kennedy, S. O'Shaughnessy, and D. Trimble, "A review of critical repeatability and reproducibility issues in powder bed fusion," *Materials and Design*, vol. 2020, p. 108346, 186.

[22] E. Malekipour and H. El-Mounayri, "Common defects and contributing parameters in powder bed fusion am process and their classification for online monitoring and control: a review," *Int J Adv Manuf Technol*, vol. 95, pp. 527–550, 2018.

[23] A. Wegner and G. Witt, "Correlation of process parameters and part properties in laser sintering using response surface modeling," *Physics Procedia*, vol. 39, pp. 480–490, 2012.

[24] S. Lee, M. Kooshkbaghi, K. Spiliotis, C. I. Siettos, and I. G. Kevrekidis, "Coarse-scale pdes from fine-scale observations via machine learning," *Chaos*, vol. 30, p. 013141, 2020.

[25] Z. Trstanova, B. Leimkuhler, and T. Lelièvre, "Local and global perspectives on diffusion maps in the analysis of molecular systems," *Proc. R. Soc. A*, vol. 476, p. 20190036, 2020.

[26] A. Ozakin and A. G. Gray, "Submanifold density estimation," *Advances in Neural Information Processing Systems*, vol. 22, pp. 1375–1382, 2009.

[27] L. McInnes and J. Healy, "Umap: Uniform manifold approximation and projection for dimension reduction," *ArXiv e-prints*, 2018.

[28] R. R. Coifman and S. Lafon, "Diffusion maps," *Appl. Comput. Harmon. Anal.*, vol. 21, pp. 5–30, 2006.

[29] R. R. Coifman, I. G. Kevrekidis, S. Lafon, M. Maggioni, and B. Nadler, "Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems," *Multiscale Model. Simul.*, vol. 7, no. 2, pp. 842–864, 2008.

[30] A. Haddad, D. Kushnir, and R. R. Coifman, "Texture separation via a reference set," *Appl. Comput. Harmon. Anal.*, vol. 36, pp. 335–347, 2014.

[31] D. Kushnir, A. Haddad, and R. R. Coifman, "Anisotropic diffusion on sub-manifolds with application to eath structure classification," *Appl. Comput. Harmon. Anal.*, vol. 32, pp. 280–294, 2012.

[32] G. Mishne and I. Cohen, "Multiscale anomaly detection using diffusion maps," *IEEE J-STSP*, vol. 7, no. 1, pp. 111–123, 2013.

[33] A. W. Long and A. L. Ferguson, "Landmark diffusion maps (l-dmaps): Accelerated manifold learning out-of-sample extension," *Appl. Comput. Harmon. Anal.*, vol. 47, no. 1, pp. 190–211, 2017.

[34] R. G. Ghanem, C. Soize, and C. A. Thimmisetty, "Optimal well-placement using probabilistic learning," *Data-Enabled Discovery and Applications*, vol. 2, no. 1, 2018.

[35] C. A. Thimmisetty, R. G. Ghanem, J. A. White, and X. Chen, "High-dimensional intrinsic interpolation using gaussian process regression and diffusion maps," *Math Geosci*, vol. 50, pp. 77–96, 2018.

[36] T. Butler, D. Estep, S. Tavener, C. Dawson, and J. J. Westerink, "A measure-theoretic computational method for inverse sensitivity problems iii: multiple quantities of interest," *SIAM/ASA J. Uncertainty Quantification*, vol. 2, pp. 174–202, 2014.

[37] D. C. Perrault-Joncas, M. Meila, and J. McQueen, "Improved graph laplacian via geometric consistency," *31st Conference on Neural Information Processing Systems*, 2017.

[38] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, Incorporated, 2015.

[39] H. W. Jones, "The recent large reduction in space launch cost," *48th International Conference on Environmental Systems*, 2018.

[40] B. Lyons, "Additive manufacturing in aerospace: Examples and research outlook," *The Bridge*, vol. 42, no. 1, pp. 13–19, 2012.