

Optimal Shape Control via L_∞ Loss for Composite Fuselage Assembly

Juan Du¹, Shanshan Cao¹, Jeffrey H. Hunt², Xiaoming Huo¹ * and Jianjun Shi¹

¹H. Milton Stewart School of Industrial and Systems Engineering,
Georgia Institute of Technology

²The Boeing Company

November 12, 2019

Abstract

Natural dimensional variabilities of incoming fuselages affect the assembly speed and quality of fuselage joins in composite fuselage assembly processes. Shape control is critical to ensure the quality of composite fuselage assembly. In current practice, the structures are adjusted to the design shape in terms of the ℓ_2 loss for further assembly without considering the existing dimensional gap between two structures. Such practice has two limitations: (1) the design shape may not be the optimal shape in terms of a pair of incoming fuselages with different incoming dimensions; (2) the maximum gap is the key concern during the fuselage assembly process, so the ℓ_∞ loss of gap after control oughts to be considered. This paper proposes an optimal shape control methodology via the ℓ_∞ loss for composite fuselage assembly process by considering the existing dimensional gap between the incoming pair of fuselages. On the other hand, due to the limitation on the number of available actuators in practice, we face an important problem of finding the best locations for the actuators among many potential locations, which makes the problem a sparse estimation problem. We are the first to solve the optimal shape control in fuselage assembly process using the ℓ_∞ model under the framework of sparse estimation, where we use the ℓ_1 penalty to control the sparsity of the resulting estimator. From statistical point of view, this can be formulated as the ℓ_∞ loss based linear regression, and under some standard assumptions, such as the restricted eigenvalue (RE) conditions, and the light tailed noise, the non-asymptotic estimation error of the ℓ_1 regularized ℓ_∞ linear model is derived to be the order of $O(\sigma\sqrt{\frac{S \log p}{n}})$, which meets the upper-bound in the existing literature. Compared to the current practice, the case study shows that our proposed method significantly reduces the maximum gap between two fuselages after shape adjustments.

Keywords: linear regression, sparsity, assembly processes, dimensional variations

*The authors gratefully acknowledge the *Strategic University Partnership between the Boeing Company and Georgia Institute of Technology*. The corresponding author is Xiaoming Huo.

1 Introduction

Recently, composite materials are widely used in large space structures due to its superior properties such as high strength-to-weight ratio. For example, an airplane of Boeing 787 comprises more than 50% composite parts by weight, and the fuselage is one key composite part of an aircraft [7]. In practice, there are natural dimensional variations in the fabrication of a fuselage due to different manufacturing batches or different suppliers [11]. When two fuselages assemble together, there is a gap, as shown in Figure 1. Such gap will influence the assembly quality and productivity of fuselage assembly. Thus, dimension variations influence the speed and the quality of composite fuselage assembly. As a result, the interface between two fuselages requires shape adjustments before the composite fuselage assembly process.

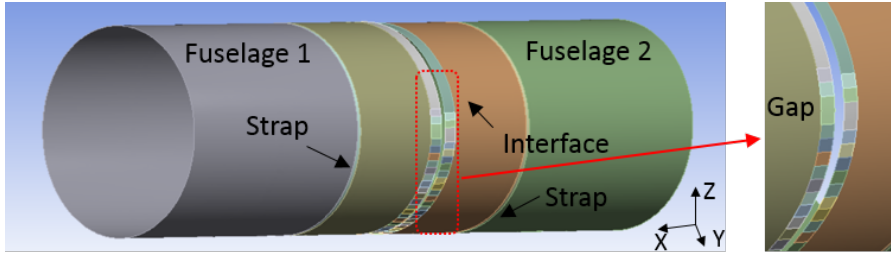


Figure 1: Illustrations of composite fuselage assembly.

In practice, actuators are used for shape adjustments of the interface between two composite fuselages, as shown in Figure 2. More details about using actuators for shape adjustments of composite fuselages can be found in [19], [20]. In this paper, we focus on the shape adjustments of interface that are close to the edge plane of the fuselage. In the state of the art, the target shape after control is the design shape, as shown in Figure 3. The dashed line indicates the design shape and the solid line shows the shape of an incoming fuselage. The arrows represent the positions of the actuators used for shape control of the interface. The current shape control strategy regards the target shape after adjustments as the design shape in terms of the ℓ_2 loss, which has two key limitations. (1) It is non-optimal. For a given pair of incoming fuselages with specific shapes, adjustment to the design shape is not optimal. In other words, there is no guarantee that the optimal shape control can be realized by shape adjustments to the design shape for two different

incoming fuselages. Here, optimality means to achieve the minimum of maximum gap on the interface between two fuselages after shape control. (2) For fuselage assembly, the maximum gap between two fuselages is the key quality indicator during the fuselage assembly process. Thus, the l_∞ loss should be considered.

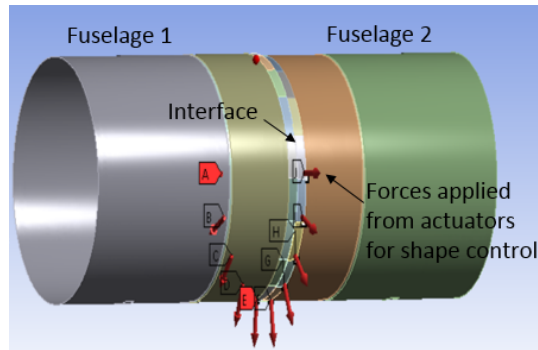


Figure 2: An illustration of shape adjustments by using actuators.

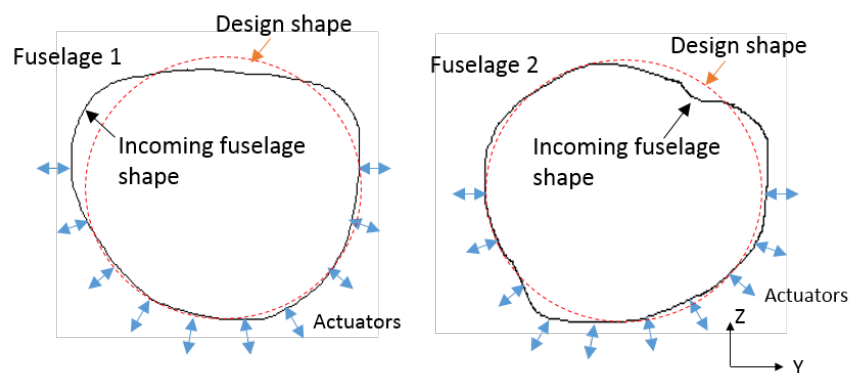


Figure 3: Schematics of interfaces between two fuselages.

In the literature, multiple efforts are made to achieve shape control for structures. To focus on the research work related to this paper, we mainly introduce the literature for composite fuselage shape control.

For the shape control of composite fuselages, Wen et al. [19] first develop a new shape control system based on the Finite Element Analysis (FEA) to improve the dimensional quality

and productivity. Also, they show the feasibility of shape control for composite fuselages by using the proposed FEA platform. Based on this FEA platform, Yue et al. [20] propose a surrogate model-based control strategy by considering the uncertainties for composite fuselage assembly. By minimizing the ℓ_2 loss of dimensional errors in relative to the design shape, they calculate the optimal forces applied from actuators for single fuselage shape control. Based on the surrogate model, Du et al. [10] propose an optimal actuator placement strategy for the shape adjustment of composite fuselages; their strategy reduces the forces applied from actuators and also dimensional deviations after shape control. However, all these works only consider shape adjustments of a single fuselage, and aim to adjust the incoming fuselages to the design shapes in terms of the ℓ_2 loss, which meets the requirement of a single shape adjustment. However, when two fuselages assemble together, the maximum gap, i.e., the ℓ_∞ norm of the gap, is one of the most important quality indicator in a fuselage assembly process. In terms of the maximum gap reduction for composite fuselage assembly, adjusting incoming fuselages to design shapes cannot guarantee the optimality given a specific pair of fuselages with different dimensional variations. Hence, a better shape control strategy is needed in the composite fuselage assembly process.

To fill the aforementioned gap, this paper proposes an optimal shape control strategy for composite fuselage assembly. We consider minimizing the maximum of dimensional gap between a pair of incoming fuselages. Instead of adjusting each fuselage into the design shape, we consider the initial gap between the pair of fuselages and optimize the adjustment into an intermediate shape.

As shown in [10], one other direct result is that the optimal placement of actuators can also lead to the improvement of the performance in shape control. In the current paper, as a part of the optimal shape control strategy for the composite fuselage assembly, we also consider the optimal placement of actuators for two fuselage assembly. Given the dimensions of the pair of incoming fuselages, a sparse learning model is proposed to link actuator forces with the weighted maximum gap deviation of a pair of fuselages. From a statistical point of view, the proposed sparse learning model is the ℓ_1 sparsity penalized ℓ_∞ loss linear regression. We also contribute to analyze the properties of the resulting estimator under a light-tailed sub-Gaussian errors assumption. Specifically, we provide the non-asymptotic upper-bound of the estimation error and prediction error in the sparse estimation problem. An alternating direction method

of multipliers (ADMM)-based algorithm is derived to solve the optimization problem. The nonzero components of the force vector correspond to the optimal actuator locations. Often in practice, the last step is to refit the model with only the selected locations and get a more accurate estimation on the unknown forces. Consequently, the optimal forces applied to the pair of fuselages can be obtained by minimizing the weighted maximum gap deviation. Finally, case studies of composite fuselage assembly process are used to verify the effectiveness of our proposed method.

1.1 Notations

Throughout this work, we will need the following notations. We will denote vectors/matrix as Y , X , β , ϵ , etc, and all the vectors are the column vectors by default. For a matrix $A \in \mathbb{R}^{m \times n}$, A_{ij} denotes the element of A at the (i, j) th location, and A_i denotes the i th column of A . Similarly, for a vector $\beta \in \mathbb{R}^m$, β_i denotes the i th element of β . For a subset $S \subset \{1, \dots, m\}$, A_S denotes sub-matrix of A containing the columns whose indices are in S , and β_S denotes a sub-vector of β that contains the elements whose indices are in S . We will use $\|\cdot\|_p$ to denote the ℓ_p norm. Specifically, for a vector $x \in \mathbb{R}^n$, the ℓ_p norm of x is defined as: $\|x\|_p := (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$. According to this definition, the ℓ_∞ norm of x is simply $\|x\|_\infty = \max_{i=1}^n |x_i|$. In the case of the Euclidean norm, which is also known as the ℓ_2 norm, we will simply omit the subscript p ($p = 2$), i.e., $\|\cdot\|$. We will denote the entrywise max norm of a matrix $A \in \mathbb{R}^{m \times n}$ as $\|A\|_\infty = \max_{1 \leq i \leq m, 1 \leq j \leq n} |A_{ij}|$. Throughout this work, we use the $\text{sign}(\cdot)$ function to indicate the sign of a vector or scalar. For $x \in \mathbb{R}$, $\text{sign}(x) = 1$ if $x > 0$ and $\text{sign}(x) = -1$ otherwise. For a vector $\beta \in \mathbb{R}^m$, $\text{sign}(\beta)$ is a vector with the i th element equal to $\text{sign}(\beta_i)$. For a vector $\beta \in \mathbb{R}^m$, the support of β is defined as $\text{supp}(\beta) = \{i : \beta_i \neq 0\}$. And $|\text{supp}(\beta)|$ denote the cardinality of the set $\text{supp}(\beta)$.

1.2 Outline

In this work, we first provide a detailed description of the physical model that is of concern, including our justifications of using a linear model with the ℓ_∞ loss function in Section 2. Then, we investigate the theoretical properties of the proposed method, and provide main theoretical results in Section 3. In Section 4, we propose an efficient algorithm to solve the proposed ℓ_∞ optimization problem. Our method is based on ADMM. A case study on fuselage assembly

process using the FEA generated data is studied in Section 5, which validates the effectiveness of our proposed approach in this type of engineering problems. Finally, we conclude our work in Section 6. All the proofs of our main theorems are provided in the Appendix.

2 Physical model

In this section, before going to the statistical analysis, we provide necessary background of our physical model. For the composite fuselage shape control, only elastic deformation is allowed during the shape adjustment. In this paper, we assume linear mechanical behavior of fuselage deformation that corresponds to the actuator forces. Consequently, the adjusted shape deviations can be formulated as

$$\delta_i = \psi_i + U_i F_i, \quad i = 1, 2. \quad (1)$$

where i indicates the i th fuselage, $\delta_i \in \mathbb{R}_i^{2n}$ is the error in Y and Z directions after shape control and n denotes the number of measurement points for each fuselage; ψ_i and $U_i \in \mathbb{R}^{2n \times m_i}$ represent dimension deviations and displacement matrix of incoming fuselage i , respectively, and m_i denotes the number of feasible positions (e.g., candidate positions where an actuator may be placed) for actuators in the i th fuselage edge plane; F_i is the applied force during shape control of the fuselage i . The physical interpretation of the displacement matrix U_i is that U_i is the deformation of all the measurement points correspond to the unit force on the structure. Without loss of generality, we assume that the locations of measurement points are the same. Therefore, the dimensional gap between two fuselages after shape adjustment can be written as

$$\Delta = \delta_2 - \delta_1 = \psi_2 + U_2 F_2 - (\psi_1 + U_1 F_1). \quad (2)$$

Similar to [10], $U_i, i = 1, 2$, can be obtained from the surrogate model. Notably, the registration between two fuselages is needed if the number of measurement points for the pair of fuselages is not the same, which can be easily solved through a method that is proposed in [8].

For the composite fuselage assembly process, the main concern before assembly is the maximum gap and its location along the interface between the two fuselages. In this paper, our objective is to minimize the weighted maximum gap between the pair of fuselages in both Y and

Z directions, i.e., Δ_{max} , which is defined as

$$\Delta_{max} = \|B\Delta\|_{\infty}, \quad (3)$$

where $B \in \mathbb{R}^{n \times n}$ is a diagonal matrix of weights, which represent the importance of the gaps in different measurement points. Such weight matrix is determined based on some engineering knowledge. For example, if we emphasize the dimensional gap on the upper fuselage, we can add more weight in the part of the B matrix that correspond to the upper fuselage.

Usually in practice, we only have a limited number of actuators available. Thus, among many potential choices of positions of actuators, we want to find the most effective ones, which may correspond to a sparse solution. Recall that the optimal actuator locations are the nonzero locations in the solution. Let m_1 and m_2 denote the numbers of all feasible locations for actuators along the pair of fuselages, while only a total number of M actuators are available for the shape control process in practice. In order to encourage the sparsity in the resulting solution, we add a sparsity-induced penalty, namely, the ℓ_1 penalty, on the unknown forces at different locations:

$$\min_{F_1, F_2} J := \|B\Delta\|_{\infty} + \lambda\|F_1\|_1 + \lambda\|F_2\|_1, \quad (4)$$

where λ is the tuning parameter, which controls the sparsity of the resulting estimation.

From the solution of Problem (4), the optimal actuator placement is obtained from the support of F_1 and F_2 for the shape adjustments of the pair of fuselages. In order to reduce the bias on the estimation of the forces, which are denoted by F_1 and F_2 , we refit the model using the optimal positions and the optimal forces can be obtained by solving the following problem:

$$\min_{F_1, F_2} J := \|B(\psi_2 - \psi_1 + (U_2)_{S_2}(F_2)_{S_2} - (U_1)_{S_1}(F_1)_{S_1})\|_{\infty} \quad (5)$$

where $S_1 = \text{supp}(F_1)$, $S_2 = \text{supp}(F_2)$.

In order to make the above physical model more evident to the statisticians without engineering background, we rewrite the physical model as follows:

$$\min_{\beta} \|Y - X\beta\|_{\infty} + \lambda\|\beta\|_1, \quad (6)$$

where $Y = B(\psi_2 - \psi_1)$, $X = [BU_1, -BU_2]$, and $\beta = [F_1', F_2']'$. More details of rewriting the physical model as a statistical problem will be provided in Section 3.

From the statistical perspective, the above formulations can be viewed as the linear regression problem with ℓ_∞ loss and ℓ_1 regularization, and then refitting the model using only the nonzero locations, which helps reduce the bias on the estimation of the forces, i.e., F_1 and F_2 . In this way, we can have a better estimation in practice.

In the statistical literature, the parameter estimation based on the ℓ_∞ loss has been studied since 1980, such as applications in the physical and environmental sciences [13, 21, 17], signal processing and systems engineering [16, 1, 6], and so on. However, the theoretical guarantee is very limited, especially in the sparse estimation framework.

There are two main directions studying regression analysis in the statistical literature. The first one focuses on the asymptotic distribution in linear regression in a well-posed setting, where we have many more observations than the number of feature variables, and there is no assumption on the sparsity of the unknown signals. Reference [15] studies the problem of ℓ_∞ (or Chebyshev) estimator, which minimizes the maximum absolute residual under the condition where the noise distribution is known to have bounded support or light tails. They derived the asymptotic distribution of the estimator in the low dimension setting.

The second one focuses on problems in the ill-posed settings, or settings where the unknown signal is sparse. In other words, there are far more feature variables than the number of observations or there are a lot of zero components in the coefficients. The Dantzig selector proposed in [5] solves the following problem:

$$\begin{aligned} \min_{\beta \in \mathbb{R}^p} \quad & \|\beta\|_1, \\ \text{s.t.} \quad & \|X^T(Y - X\beta)\|_\infty \leq \lambda\sigma, \end{aligned} \tag{7}$$

where the constrained optimization problem seeks to minimize the ℓ_1 sparsity objective function within the feasible region, where $\|X^T(Y - X\beta)\|_\infty \leq \lambda\sigma$. It can be easily seen that the term $X^T(Y - X\beta)$ is simply the first order derivative of the least square loss. Thus, the constraint $\|X^T(Y - X\beta)\|_\infty \leq \lambda\sigma$ will ensure that the scale of the first order derivative of the least square loss is very small. Essentially, Dantzig selector solves a similar problem as the lasso, i.e., minimizing the least square loss [14].

However, theories to the previous models do not apply to our problem, where we have the incentive to minimize $\|Y - X\beta\|_\infty$ loss. In Section 3, we rewrite the above sparse estimation problem using the ℓ_∞ loss in the statistical language in details, and provide theoretical results on

the non-asymptotic upper-bounds of the estimation error and prediction error under the linear model setting with light tailed sub-Gaussian errors.

3 Theoretical Results

In this section, we rewrite the above physics model in statistical language to make the problem clearer with reader with no engineering background. Specifically, we give the statistical formulation in Subsection 3.1. Then we present our main theoretical results in Subsection 3.2.

3.1 Model in statistical language

We will consider the linear model throughout this work:

$$Y = X^T \beta + \epsilon. \quad (8)$$

For all observations, it is assumed that all elements in the vector ϵ , i.e., ϵ_i , for $i = 1, \dots, n$, are independent sub-Gaussian random variables with common variance σ^2 and mean 0. As in our motivated example, we want to estimate the forces, which will minimize the ℓ_∞ norm of the gap between fuselages along the interface. Thus, we propose the following formulation of the estimation problem:

$$\min_{\beta} \|Y - X\beta\|_\infty, \quad (9)$$

where $Y - X\beta$ is the estimated gap between two fuselages after control. In our motivated application, we believe the underlying forces are sparse to find the optimal locations of the actuators, i.e., the β vector is sparse. Thus, in this work, we mainly focus on the scenario where the β vector is sparse. However, our results also applies for the high-dimensional scenario where $p \gg n$, and the signal (i.e., the coefficient vector) is sparse to guarantee the feasibility of the system. In order to derive the sparsity of the resulting solution, we will consider the regularized version of Problem (9) with ℓ_1 penalty on the unknown parameter, which is the Problem (6). Notably, the Problem (6) is equivalent to the formulation as follows [18],

$$\begin{aligned} \min_{\beta \in \mathbb{R}^p} \quad & \|\beta\|_1, \\ \text{s.t.} \quad & \frac{1}{\sqrt{n}} \|Y - X\beta\|_\infty \leq \lambda_0, \end{aligned} \quad (10)$$

where $\lambda_0 > 0$ is a tuning parameter to control the sparsity of the resulting estimation. Also, λ_0 and λ has the one-to-one correspondence. It can be easily checked that Problem (10) is a convex system and can be recast as a linear programming (LP) problem:

$$\begin{aligned} \min_{\gamma \in \mathbb{R}^{+p}} \quad & \sum_{i=1}^p \gamma_i \\ \text{s.t.} \quad & -\gamma_i \leq \beta_i \leq \gamma_i, \text{ for } i = 1, \dots, p \\ & -\lambda_0 \leq \frac{1}{\sqrt{n}}(Y - X\beta)_j \leq \lambda_0, \text{ for } j = 1, \dots, n. \end{aligned} \tag{11}$$

For the interest of our proof, we use the constrained version as shown in Problem (10) in our theoretical analysis.

3.2 Main results

In this section, we present our main theoretical results on the estimation error of the resulting estimator that is proposed in Formulation (10). The proofs are postponed to the Appendix.

3.2.1 Restricted strong convexity assumption

In this subsection, we will summarize the assumptions we need for deriving the main results. These are standard assumptions used in the statistical literature on recovering the sparse signals.

Definition 3.1. *The restricted strong convexity (RSC) condition on model matrix X with respect to a set \mathcal{C} is the following, there exists some constant $\gamma > 0$ such that:*

$$\frac{\frac{1}{n}\nu X^T X \nu}{\|\nu\|_2^2} \geq \gamma \text{ for all nonzero } \nu \in \mathcal{C},$$

where γ is called the restricted eigenvalue bound with regard to \mathcal{C} .

Assumption 3.2. *The restricted strong convexity (RSC) condition holds on the following set:*

$$\mathcal{C} = \{\nu \in \mathbb{R}^p \mid \|\nu_{S^c}\|_1 \leq \|\nu_S\|_1\}.$$

We have $\mathcal{C} \subset \mathbb{R}^p$ strictly since it is of the form of a cone.

The RSC (Assumption 3.2) is a standard assumption in the literature for proving the consistency results of regularized high-dimensional sparse estimation problems. According to [4, 5],

the assumption on restricted strong convexity condition will hold with high probability for random design matrices X , such as a random matrix with i.i.d. Gaussian entries, or Rademacher entries.

3.2.2 Estimation error upper-bound

Before we give the estimation error upper-bound, we first prove a lemma, which states that the unknown ground truth β^* , is feasible to Problem (10) with high probability.

Lemma 3.3. *Let $\lambda_0 = \sigma\sqrt{\alpha\frac{\log p}{n}}$ for some $\alpha > 2$, then with probability exceeding $1 - \frac{2n}{p} \cdot (\frac{1}{p})^{\frac{\alpha-2}{2}}$, the unknown ground truth β^* is a feasible solution to Problem (10).*

This lemma basically tells that, if the noise in the data generation is sub-Gaussian, when $\frac{2n}{p} \cdot (\frac{1}{p})^{\frac{\alpha-2}{2}} \rightarrow 0$, the ground truth β^* lies in the feasible region of Problem (10). Further, if the noise has a bounded distribution within the interval $[-\lambda_0, \lambda_0]$, β^* lies in the feasible region of Problem (10) with probability 1.

Theorem 3.4. *Let $\hat{\beta}$ denote the optimal solution to Problem (10). Suppose β^* is any unknown sparse ground truth with $|\text{supp}(\beta)| \leq S$, and the restricted strong convexity assumption holds for the observed feature matrix X with γ in \mathcal{C} . Let $\lambda_0 = O(\sigma\sqrt{\frac{\log p}{n}})$. We have the following non-asymptotic upper-bound on the estimation error:*

$$\|\hat{\beta} - \beta^*\|_2 \leq O(\sigma\sqrt{S\log p}). \quad (12)$$

According to the above result, the performance of sparse parameter estimation from Problem (10) is guaranteed. Not only we will be able to recover the sparse unknown parameter in the high-dimensional scenario, but also we can bound the mean squared error (MES) of the resulting estimation in an order of $O(\sigma\sqrt{S\log p})$. This upper-bound in fact is consistent with the MSE in [5].

However, we may obtain a better bound if we add one more condition: the maximum absolute column sum of the matrix norm of the observed feature matrix X : $\|X\|_1 := \max_{i=1}^p \|X_i\|_1$ is also in the order of \sqrt{n} . This condition usually holds with sparse matrix, where there are a lot of 0 entries, or even matrix with columns, where the entry values decay exponentially. This is almost the case for the fuselage shape control problem that the feature matrix $X = (BU_1, -BU_2)$

has a lot of near zero (magnitude of 10^{-6}) entries due to the holding fixtures of the fuselages. Specifically, we state our conclusion in the following theorem:

Theorem 3.5. *Let $\hat{\beta}$ denote the optimal solution to Problem (10). Suppose β^* is any unknown sparse ground truth with $|\text{supp}(\beta)| \leq S$, and the restricted strong convexity assumption holds for the observed feature matrix X with γ in \mathcal{C} . Let $\lambda_0 = O(\sigma\sqrt{\frac{\log p}{n}})$. If we further have that $\|X\|_1 = O(\sqrt{n})$, we have the following tighter non-asymptotic upper-bound on the estimation error:*

$$\|\hat{\beta} - \beta^*\|_2 \leq O(\sigma\sqrt{\frac{S \log p}{n}}). \quad (13)$$

The proofs to Theorem 3.4 and 3.5 are in Appendix A.2 and A.3, respectively. The above Theorem 3.4 and 3.5 essentially state that with high probability, the optimal solution from our proposed formulation in (10) is very close to the unknown ground truth, with their difference bounded by $O(\sigma\sqrt{\frac{S \log p}{n}})$.

3.2.3 Prediction Error Upper-Bound

Theorem 3.6. *Under the assumptions of Theorem 3.5, we have the following upper-bound for the prediction error:*

$$\|X(\hat{\beta} - \beta^*)\|_2 \leq O(\sigma\sqrt{\frac{S \log p}{n}}). \quad (14)$$

The proof is simple, which is due to the by-product in the proof of Theorem 3.5 and is shown in Appendix A.4. This theorem tells that even if we are aiming to find a sparse solution to minimize the ℓ_∞ loss function, the resulting solution will still produce a good estimation in the sense of the mean squared error.

4 Algorithm

In this section, we proposed an algorithm to efficiently solve the optimization problem (4) and (5) based on the ADMM algorithm [3]. Compared to the gradient method that is sensitive to the choice of the step-size, the ADMM algorithm is robust to the parameter selection for algorithm convergence [12]. The ADMM algorithm is well suited to the convex optimization problem with

separate objectives, which is the case of the optimization problem (4), so we adopt ADMM algorithm for parameter estimation. Problem (4) can be further written as

$$\min_y \left(\sum_{i=1}^2 f_i(y_i) + g(z) \right), \quad (15)$$

where

$$\begin{aligned} f_1(y_1) &= \|B(\psi_2 - \psi_1 + U_2 F_2 - U_1 F_1)\|_\infty = \|y_1\|_\infty, \\ f_2(y_2) &= \lambda \|F_1\|_1 + \lambda \|F_2\|_1 = \lambda \|y_2\|_1, \\ y_2 &= [F_1', F_2']', \\ y_1 &= Ay_2 + b, A = [-BU_1, BU_2], b = B(\psi_2 - \psi_1), y = [y_1', y_2']'. \\ g(z) &= \begin{cases} \infty, & Ez \neq b, \\ 0, & \text{otherwise,} \end{cases} \\ E &= (I, -A), z = y. \end{aligned}$$

According to the ADMM formula, the algorithm for the above problem can be written as

$$y_i^{k+1} = \text{prox}_{f_i}(z_i^k - u_i^k), \quad (16)$$

$$z^{k+1} = \text{prox}_g(y^{k+1} + u^k), \quad (17)$$

$$u^{k+1} = y^{k+1} + u^k - z^{k+1}. \quad (18)$$

Here, $\text{prox}_f(x)$ is the proximal operator of f , which defined as

$$\text{prox}_f(x) = \arg \min_u (f(u) + 1/2 \|u - x\|_2^2). \quad (19)$$

The proximal operator prox_{f_i} and prox_g can be analytically derived if we formulated the optimization problem in this way. From Chapter 6 of [2], the proximal operators of functions $f_i(y_i)$ and $g(z)$ can be derived as

$$\text{prox}_{f_1}(y_1) = y_1 - P_{B_{\|\cdot\|_1}[0,1]}(y_1), \quad (20)$$

$$\text{prox}_{f_2}(y_2) = S_{\lambda/\rho}(y_2), \quad (21)$$

$$\text{prox}_g(z) = z - E'(EE')^{-1}(Ez - b), \quad (22)$$

where $P_{B_{\|\cdot\|_1}[0,1]}(\cdot)$ is the projection onto the unit ball $G = B_{\|\cdot\|_1[0,1]} = \{x \in \mathbb{R}^n : \|x_1\|_1 \leq 1\}$, and

$$P_{B_{\|\cdot\|_1}[0,1]}(y_1) = \begin{cases} y_1, & \|y_1\|_1 \leq 1, \\ S_{\lambda^*}(y_1), & \|y_1\|_1 > 1, \end{cases}$$

where λ^* is any positive root of the non-increasing function $\varphi(\lambda) = \|S_\lambda(y_1)\|_1 - 1$. S_λ is the soft thresholding operator of $\|\cdot\|_1$, which is

$$S_\lambda(v) = (v - \lambda)_+ - (-v - \lambda)_+, \quad (23)$$

where $(x)_+$ is short for $\max\{x, 0\}$.

Due to the convexity of Problem (4), the convergence of the ADMM is guaranteed given a reference point and we can obtain the global optimum of the optimization Problem (4). In practice, Boyd et al. [3] suggested that a reasonable termination of the ADMM is that both the primal and dual residuals should be small, i.e.,

$$\|r^k\|_2 \leq e_1, \|s^k\|_2 \leq e_2, \quad (24)$$

where $r^k = y^k - z^k$ is the primal residual, and $s^k = -\rho(z^k - z^{k-1})$ is the dual residual. $e_1 > 0$ and $e_2 > 0$ are the tolerances, which can be chosen by using an absolute criterion and a relative criterion, i.e., we have

$$e_1 = \sqrt{m}e_3 + e_4 \max\{\|z^k\|_2, \|y^k\|_2\}, \quad (25)$$

$$e_2 = \sqrt{m}e_3 + e_4 \|\rho u^k\|_2, \quad (26)$$

where $m = m_1 + m_2$, $e_3 > 0$ and $e_4 > 0$ are an absolute tolerance and a relative tolerance, respectively. The following algorithm is proposed for estimations of F_1 and F_2 , and K is the maximum number of iterations.

Since we have the sparsity requirement of $\|F_1\|_0 + \|F_2\|_0 \leq M$, the λ value needs some investigations. Binary search can be used for computing for this value to meet the aforementioned requirement. The lower bound of the search range is trivial, which is 0, but the upper bound needs some theoretical investigations. In the following Proposition 4.1, we provide an upper bound of the search range.

Proposition 4.1. *If we choose the tuning parameter λ in the Problem (4) such that $\lambda > \|X\|_\infty$, $X = [BU_1, -BU_2]$, there exists one unique global minimum to (4), with $\hat{\beta} = [F'_1, F'_2]' = 0$.*

1: Input: $\rho, \lambda, e_3, e_4, m_1, m_2, n, B, U_1, \psi_1, U_2, \psi_2, A, b$
2: Initialize z^0, u^0
3: *loop*:
4: **for** $k = 1 : K$ **do**
5: $u_1^k = \{u_i^k\}_{i=1, \dots, 2n}, u_2^k = \{u_i^k\}_{i=1+2n, \dots, 2n+m_1+m_2},$
6: $z_1^k = \{z_i^k\}_{i=1, \dots, 2n}, z_2^k = \{z_i^k\}_{i=1+2n, \dots, 2n+m_1+m_2},$
7: $y_1^{k+1} = z_1^k - u_1^k - P_{B_{\|\cdot\|_1[0,1]}}(z_1^k - u_1^k),$
8: $y_2^{k+1} = S_{\lambda/\rho}(z_2^k - u_2^k),$
9: $y^{k+1} = [(y_1^{k+1})', (y_2^{k+1})']',$
10: $z^{k+1} = y^{k+1} + u^k - E'(EE')^{-1}(Ez - b),$
11: $u^{k+1} = u^k + y^{k+1} - z^{k+1},$
12: $r^{k+1} = y^{k+1} - z^{k+1},$
13: $s^{k+1} = \rho(z^{k+1} - z^k),$
14: $e_1 = \sqrt{m}e_3 + e_4 \max\{\|z^{k+1}\|_2, \|y^{k+1}\|_2\},$
15: $e_2 = \sqrt{m}e_3 + e_4 \|\rho u^{k+1}\|_2,$
16: **if** $\|r^{k+1}\|_2 \leq e_1,$ and $\|s^{k+1}\|_2 \leq e_2$ **then**
17: **return** y_2
18: **end if**
19: **end for**

Algorithm 1: The ADMM of estimating $y_2 = [F'_1, F'_2]'$

The main idea to prove this proposition is that at any point that is not the origin, we can find a direction, along which, the objective value decreases. Due to space limit, we postpone the proof into Appendix A.5.

According to Proposition 4.1, we can set $\|X\|_\infty$ as the upper bound of the search range to find the λ value that meets the requirement $\|F_1\|_0 + \|F_2\|_0 = M$. The nonzero components of F_1 and F_2 correspond to the optimal locations of actuators for shape control of fuselage 1 and fuselage 2, respectively. Then, the forces from each actuator can be found by solving the optimization problem (5). Similarly, the ADMM algorithm can also be used in (5) to get the optimal solution by the Algorithm 2 in Appendix B.

5 Case study

In this section, we use the fuselage assembly process to validate the proposed methodology. We present our detailed experiment setting and report our results in Subsections 5.1 and 5.2, respectively.

5.1 Numerical setting

In this case study, we use an FEA model [19] to generate data and validate the proposed methodology. The FEA model used in this case study has been validated with the experimental data, and more details about FEA model can be found in [19, 20]. For shape adjustments of single fuselage, the method from [10] achieves the best control performance in terms of reduction of shape deviations after adjustments and the forces applied during shape adjustments.

In order to make a fair comparison between our methods and the best result of composite fuselage shape control in current literature, we use the same FEA data from [10]. There are 20 incoming fuselages with different dimensional variations. Based on engineering practice, in each pair of the fuselages, $m_1 = m_2 = 18$ feasible actuators are equally placed from -12 degrees to 192 degrees in the lower part of each fuselage. More details about the data generation can be found in [10]. $M = M_1 + M_2 = 18$ actuators are used for shape adjustments of two fuselages in this case study. The number of measurement points along the interface of the pair of fuselages are $n_1 = n_2 = n = 182$, weight matrix is set to be $B = \text{diag}(1/n)$, and $U_1 = U_2$ are chosen to be

the same with [10] for comparison purpose.

In practice, some elements of displacement matrix $U_i, i = 1, 2$, are very small due to the structures of composite fuselages, such as the elements of U_i near the fixture. Specifically, we checked our displacement matrix U_i and find that the element scales decays almost exponentially for each column, which satisfies our column norm assumption in Theorem 3.5. To achieve a better computational performance, such as reducing floating point errors [9], we multiply a large constant number L_N in the objective function $\|B(\psi_2 + U_2 F_2 - (\psi_1 + U_1 F_1))\|_\infty$ and $\|B(\psi_2 + U_2^c F_2^c - (\psi_1 + U_1^c F_1^c))\|_\infty$ in the optimization Problems (4) and (5), which does not change the optimal solution. In this way, the computational problems induced by U_i matrix can be alleviated in real implementations. In this case study, we set $L_N = 10^7$. In terms of the ADMM algorithm, we set $\rho = 1, e_3 = 10^{-6}, e_4 = 10^{-5}$.

5.2 Results of the proposed method

In this case study, we randomly pick up two different fuselages from 20 incoming fuselages and have 50 replications. Consequently, we have 50 pairs of fuselages for assembly. According to the engineering practice, we use root mean square gap (RMSG), maximum gap (MG), and maximum force (MF) to evaluate the control results. These quantities are defined as follows:

$$RMSG := \frac{1}{n} \sqrt{\Delta' \cdot \Delta}, \quad (27)$$

$$MG := \sqrt{\|\Delta_Y * \Delta_Y + \Delta_Z * \Delta_Z\|_\infty}, \quad (28)$$

$$MF_i := \|F_i\|_\infty, i = 1, 2, \quad (29)$$

where ‘*’ indicates the Hadamard product, and $\Delta_Y \in \mathbb{R}^n$ and $\Delta_Z \in \mathbb{R}^n$ are the first and last n elements of Δ , respectively. Thus, MG captures the maximum gap for two fuselage assembly. Notably, we only evaluate the MF_i for the safety purpose during the fuselage shape control, since large force may destroy the fuselage structure during fuselage shape control. The control results of the proposed method on these 50 pairs of fuselages assembly are shown in Table 1. Since we care more about the maximum gap after control and maximum force used for shape control in

fuselage assembly process, we also listed the maximum (max) of RMSG, MG and MF for these 50 pairs. The results show that the proposed method performs well in terms of the maximum gap by using relatively smaller forces, which is acceptable for fuselage assembly in practice. The following Figure 4 and Figure 5 show the box-plots of the fuselage gap after control and the maximum forces used for shape control in these 50 pairs.

	RMG (inches)	MG (inches)	MF_1 (lbs)	MF_2 (lbs)
Mean	0.0014	0.0034	304.73	307.42
Max	0.0023	0.0060	553.83	776.86
Std	0.0004	0.0011	101.43	131.01

Table 1: Control results of our method on 50 pairs of fuselages.

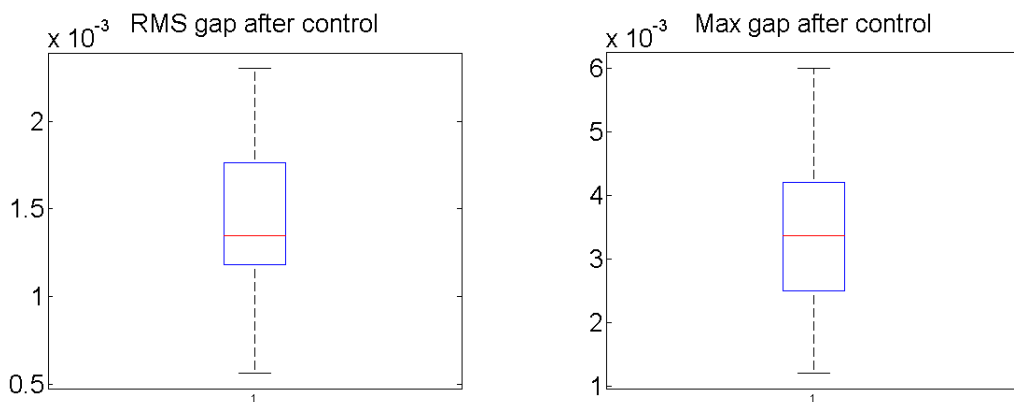


Figure 4: The boxplots of RMS gap and Max gap of 50 pairs of fuselages after control by the proposed method.

5.2.1 Comparisons

In this subsection, we compare the proposed method with the current practice [10], which is to force the incoming fuselage into the design shape in terms of the ℓ_2 loss, and then assemble. We

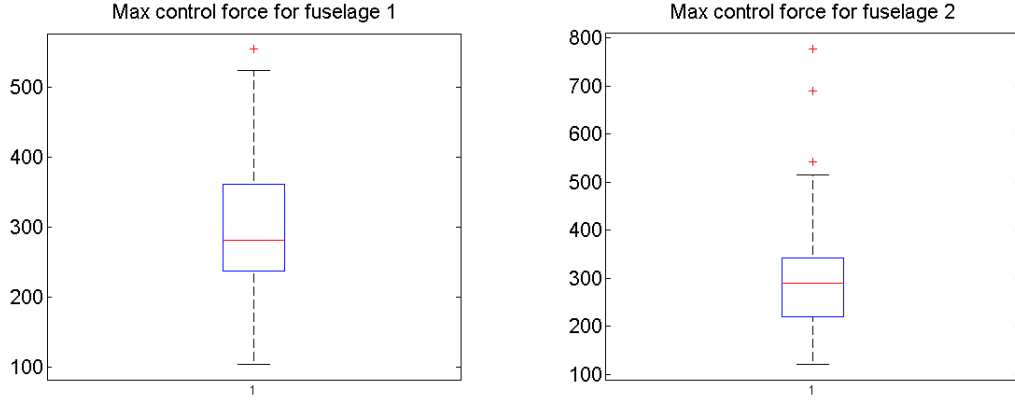


Figure 5: The boxplots of max control forces for fuselage 1 and fuselage 2 by the proposed method.

randomly choose one pair of fuselages to compare the performance of the proposed method with the one in [10]. Figure 6 shows the incoming fuselage dimensions at the interface between two fuselages and the proposed acutator locations of fuselage 1 and fuselage 2. Similarly, we also show the incoming fuselage dimensions and the acutator locations from [10] of fuselage 1 and fuselage 2 in Figure 7. To show the incoming dimensional variations more clearly, we magnify the distortions 150 times since the dimensional error is much smaller than the dimension of the fuselage. The marked round points are the locations of actuators. As shown in the Figure 6 and Figure 7, the optimal actuator locations of the proposed method are different from [10]. This is because our method aims to minimize the maximum gap between the pair of assembly fuselages, while [10] aims to control single fuselage into design shape and then assembly. We also show the shape adjustments of the proposed method and the one in [10] for this pair of fuselages in Figure 8 and Figure 9, respectively. The black dots show the intial dimensional gap between two fuselages in Y and Z directions; the green dots and blue dots show the adjustment from fuselage 1 and fuselage 2 in Y and Z directions; the red dots are the final adjusted gap after control in Y and Z directions. As shown in Figure 8, the shape adjustments of our method reduces the intial gap and successfully adjusts the shape. Compared to our method, the method from [10] adjusts the shape of incoming fuselage seperately (shown in Figure 9) since the goal of this method is to

adjust the shape of the fuselage into the design shape. The control performance is not as good as the proposed method since we can see larger variations around 0 in Figure 9. This pair of fuselages illustrates the effectiveness of the proposed method in detail.

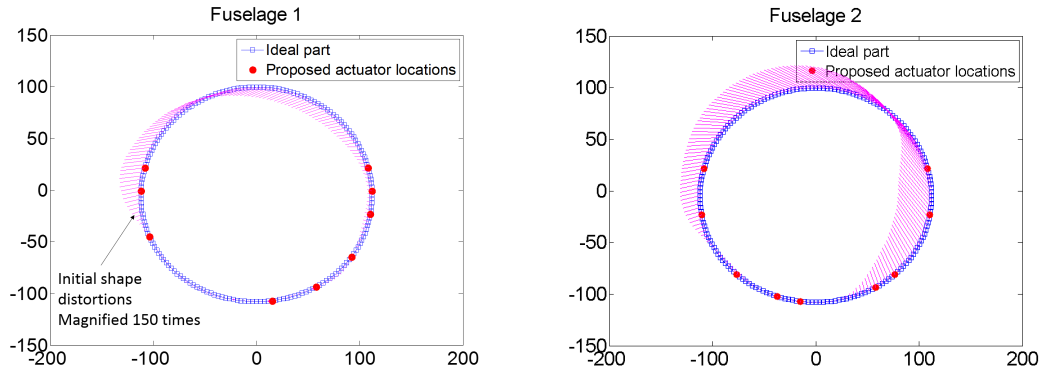


Figure 6: Incoming fuselage dimensions and the proposed actuator locations for fuselage 1 and fuselage 2.

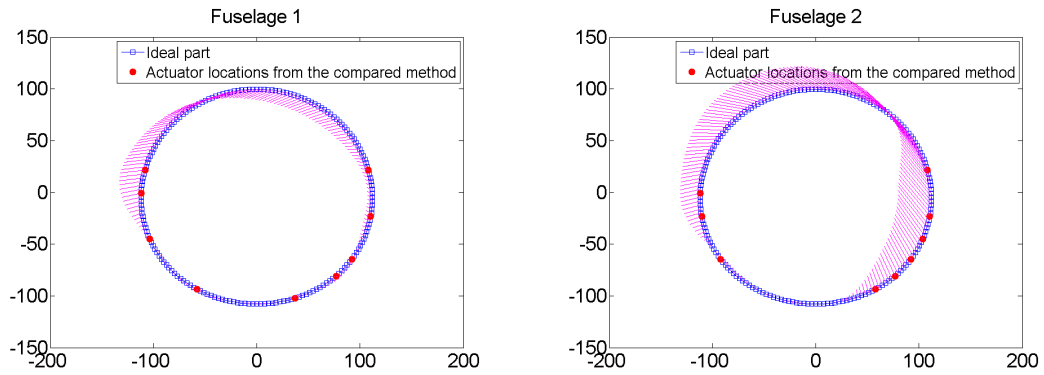


Figure 7: Incoming fuselage dimensions and the actuator locations from [10] for fuselage 1 and fuselage 2.

In order to better illustrate the improvement of the proposed method, we calculated the improvements on the gap after shape adjustments, and use statistical t test to test the significance of the improvement on the gap. The improvement results on mean and standard deviation (std)

of the gap after adjustments are listed in Table 2. Notably, the improvement means the control results via current practice [10] minus the results of the proposed method. The null hypothesis H_0 : the improvement comes from a distribution with mean zero, i.e., no significant improvement. The alternative hypothesis H_1 : the mean of improvement is greater than 0 (right-tailed test), which indicates the proposed method reduces the dimensional gap after control. The p -value is also listed in Table 2. As shown in Table 2, compared with [10], the proposed method significantly reduces the maximum gap (MG) and RMS gap (RMSG) for two fuselage assembly. The box-plots of MG and RMSG improvement are shown in Figure 10.

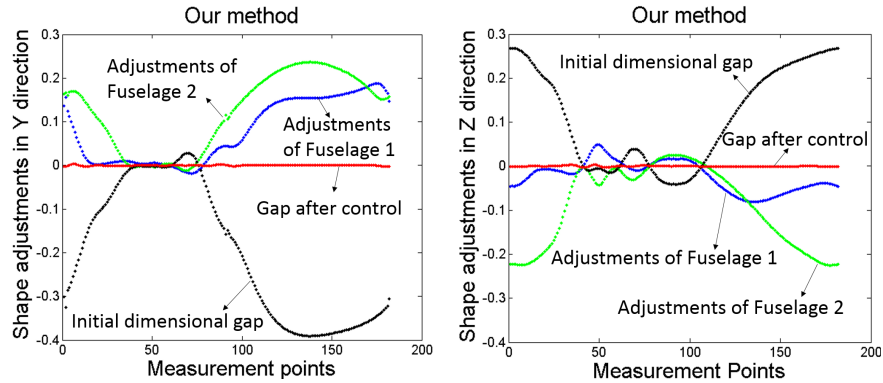


Figure 8: Incoming fuselage dimensions and the actuator locations from [10] for fuselage 1 and fuselage 2.

MG Improvement (inches)			RMSG Improvement (inches)		
Mean	Std	P-value	Mean	Std	P-value
0.0154	0.0074	6.19×10^{-20}	0.0048	0.0023	5.40×10^{-20}

Table 2: Gap reduction of our method compared with method from [10].

5.3 Discussions

A contribution in this paper is that we propose an optimal shape control strategy for composite fuselage assembly. In practice, the main concern during the composite fuselage assembly process

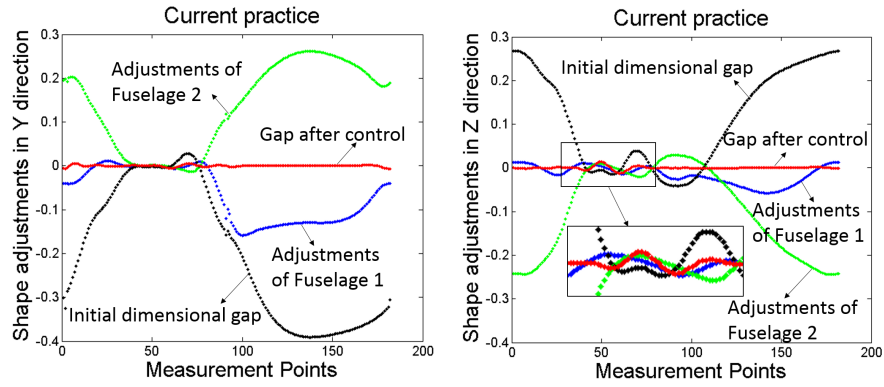


Figure 9: Incoming fuselage dimensions and the actuator locations from [10] for fuselage 1 and fuselage 2.

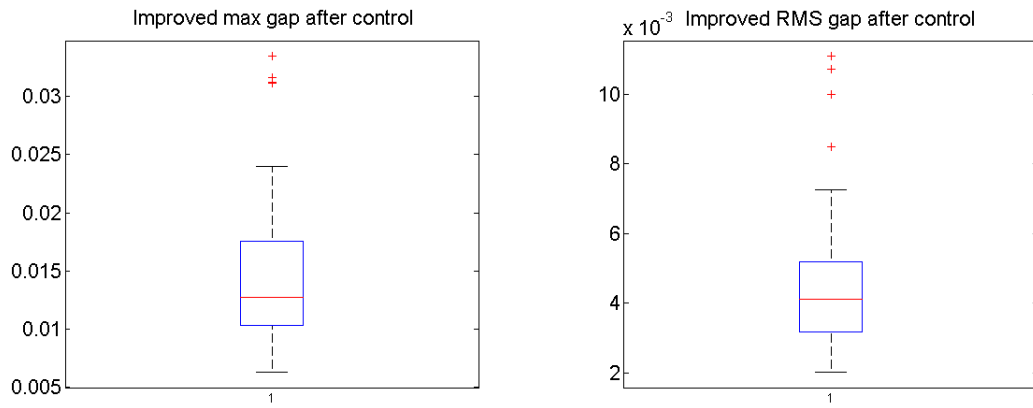


Figure 10: The boxplots of improved max and RMS gap after control compared with method from [10].

is the maximum of the dimensional gap. Consequently, the objective of this paper is to minimize the maximum gap in composite fuselage shape control instead of using the square error as in [10]. The goal of [10] is to achieve the optimal actuator placement for shape control of single fuselage, and the target shape after shape control is the design shape. In the case study, our method performs much better than the method in [10] in terms of the maximum gap after the shape control. Also, by considering the initial gap between the pair of fuselages, the resulting RMSG is also smaller than those in [10].

6 Conclusions

This paper proposes an optimal shape control strategy for composite fuselage assembly. Due to natural dimensional variations of fuselages, there is a gap on the interface of two fuselages before assembly. The current practice adjusted the shape of each fuselage to the design shape in terms of the ℓ_2 loss and then assemble, which is not optimal for the case of two fuselage assembly. Our contribution is to consider the initial gap of the pair of incoming fuselages and propose a sparse learning model, which aims to minimize the maximum gap (ℓ_∞ loss) after shape control. From the statistical view, the proposed model is ℓ_1 sparsity-driven penalized ℓ_∞ loss linear regression. Under the linear model with light tailed sub-Gaussian errors assumption, we provide the non-asymptotic upper-bound of the estimation error, measurement error, and prediction error. The real case studies of composite fuselage assembly validate the proposed method according to its control performance. Compared with the current literature, the case study shows that our method achieves significant reduction of the maximum gap after control. Notably, although our method is demonstrated for optimal shape control in composite fuselage assembly process, the methodology can be extended for optimal shape control of assembly process in other structures.

References

- [1] Alin Alecu, Adrian Munteanu, Jan PH Cornelis, and Peter Schelkens. Wavelet-based scalable l-infinity-oriented compression. *IEEE Transactions on Image Processing*, 15(9):2499–2512, 2006.

- [2] Amir Beck. *First-order methods in optimization*, volume 25. SIAM, 2017.
- [3] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, Jonathan Eckstein, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine learning*, 3(1):1–122, 2011.
- [4] Emmanuel Candes and Terence Tao. Decoding by linear programming. *arXiv preprint math/0502327*, 2005.
- [5] Emmanuel Candes, Terence Tao, et al. The dantzig selector: Statistical estimation when p is much larger than n . *The annals of Statistics*, 35(6):2313–2351, 2007.
- [6] E. Castillo, C. Castillo, and A.S. et al. Hadi. Combined regression models. *Computational Statistics*, 24:37–66, 2009.
- [7] Krishan K Chawla. *Composite materials: science and engineering*. Springer Science & Business Media, 2012.
- [8] Haili Chui and Anand Rangarajan. A new algorithm for non-rigid point matching. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662)*, volume 2, pages 44–51. IEEE, 2000.
- [9] James W Demmel. *Applied numerical linear algebra*, volume 56. Siam, 1997.
- [10] J. Du, X. Yue, J. H. Hunt, and J. Shi. Optimal placement of actuators using sparse learning for composite fuselage shape control. *ASME Transactions, Journal of Manufacturing Science and Engineering*, 2019.
- [11] Dominic Gates. Boeing finds 787 pieces aren't quite a perfect fit. *Rapport technique, Seattle Times*, 2007.
- [12] Euhanna Ghadimi, André Teixeira, Iman Shames, and Mikael Johansson. Optimal parameter selection for the alternating direction method of multipliers (ADMM): quadratic problems. *IEEE Transactions on Automatic Control*, 60(3):644–658, 2015.
- [13] F James. Fitting tracks in wire chambers using the chebyshev norm instead of least squares. *Nuclear Instruments and Methods in Physics Research*, 211(1):145–152, 1983.

- [14] Gareth M James, Peter Radchenko, and Jinchi Lv. Dasso: connections between the dantzig selector and lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(1):127–142, 2009.
- [15] KEITH KNIGHT. On the asymptotic distribution of the ℓ_∞ estimator in linear regression. 2017, manuscript.
- [16] Mario Milanese and Gustavo Belforte. Estimation theory and uncertainty intervals evaluation in presence of unknown but bounded errors: Linear families of models and estimators. *IEEE Transactions on automatic control*, 27(2):408–414, 1982.
- [17] Chong Qi. Theoretical uncertainties of the duffo–zucker shell-model mass formulae. *Journal of Physics G: Nuclear and Particle Physics*, 42(4):045104, 2015.
- [18] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.
- [19] Yuchen Wen, Xiaowei Yue, Jeffrey H Hunt, and Jianjun Shi. Feasibility analysis of composite fuselage shape control via finite element analysis. *Journal of manufacturing systems*, 46:272–281, 2018.
- [20] Xiaowei Yue, Yuchen Wen, Jeffrey H Hunt, and Jianjun Shi. Surrogate model-based control considering uncertainties for composite fuselage assembly. *Journal of Manufacturing Science and Engineering*, 140(4):041017, 2018.
- [21] Ali Zolghadri and David Henry. Minimax statistical models for air pollution time series. application to ozone time series data measured in bordeaux. *Environmental Monitoring and Assessment*, 98(1-3):275–294, 2004.

Appendix

All the proofs in this paper are provided in this appendix. The proofs of main results in Section 3.2 are in Appendix A, and the Algorithm 2 in Section 4 is provided in Appendix B. Specifically, the proofs of Lemma 3.3, Theorem 3.4 and Theorem 3.5 are given in A.1, A.2 and A.3, respectively.

A Proofs in Section 3.2

Throughout this proof, we will assume that the variance parameter in Formulation (8): $\sigma^2 = 1$. We will also assume that the observed feature matrix X is standardized with column mean as 0 and standard deviation in the order of \sqrt{n} . The general case would follow from a simple rescaling procedure. Denote $\mathcal{S}_0 := \text{supp}(\beta^*)$, the support set of the ground truth β^* .

A.1 Proof of Lemma 3.3

Proof. Given the data generation mechanism in (8), the random error vector ϵ is independent sub-Gaussian random variables. We thus have the following tail inequality:

$$\mathbb{P}\left(\frac{\epsilon_i}{\sqrt{n}} > t\right) \leq 2e^{-\frac{nt^2}{2}}.$$

By Bonferroni bound, we have

$$\mathbb{P}\left(\left\|\frac{\epsilon_i}{\sqrt{n}}\right\|_\infty > t\right) \leq 2ne^{-\frac{nt^2}{2}}.$$

Let $t = O\left(\sqrt{\frac{\log p}{n}}\right)$, say $\lambda_0 = t = \sqrt{\alpha \frac{\log p}{n}}$ for some $\alpha > 2$, we have with probability larger than $1 - \frac{2n}{p} \cdot \left(\frac{1}{p}\right)^{\frac{\alpha-2}{2}}$,

$$\left|\frac{\epsilon_i}{\sqrt{n}}\right| = \frac{1}{\sqrt{n}}|Y_i - \langle (X^T)_i, \beta^* \rangle| \leq \lambda_0,$$

for all $i = 1, \dots, n$.

□

A.2 Proof of Theorem 3.4

Proof. In order to prove the statement, we will need the following lemma.

Lemma 1.1. Denote $\nu = \hat{\beta} - \beta^*$, the difference vector of the optimal solution and the ground truth. We have $\nu \in \{x \in \mathbb{R}^p : \|x_{\mathcal{S}_0^c}\|_1 \leq \|x_{\mathcal{S}_0}\|_1\}$, which is a cone in \mathbb{R}^p .

Proof. Due to the definition of $\hat{\beta}$, which is the optimal solution to (10), and β^* is a feasible solution to (10), we have

$$\|\hat{\beta}\|_1 \leq \|\beta^*\|_1.$$

Thus, we have

$$\begin{aligned} \|\beta^*\|_1 &\geq \|\hat{\beta}\|_1 \\ &= \|\beta^* + \nu\|_1 \\ &= \|(\beta^* + \nu)_{\mathcal{S}_0}\|_1 + \|(\beta^* + \nu)_{\mathcal{S}_0^c}\|_1 \\ &\geq \|\beta_{\mathcal{S}_0}^*\|_1 - \|\nu_{\mathcal{S}_0}\|_1 + \|\nu_{\mathcal{S}_0^c}\|_1 \\ &\geq \|\beta^*\|_1 - \|\nu_{\mathcal{S}_0}\|_1 + \|\nu_{\mathcal{S}_0^c}\|_1, \end{aligned}$$

which proves the lemma. □

According to the restricted eigenvalue assumption, we have the following inequality:

$$\gamma \|\nu\|_2^2 \leq \frac{1}{n} \|X\nu\|_2^2. \quad (30)$$

On the other hand,

$$\begin{aligned} \frac{1}{n} \|X\nu\|_2^2 &= \frac{1}{n} \nu^T X^T X \nu \\ &\leq \frac{1}{n} \|X^T X \nu\|_\infty \|\nu\|_1. \end{aligned} \quad (31)$$

Since $\|\nu_{\mathcal{S}_0^c}\|_1 \leq \|\nu_{\mathcal{S}_0}\|_1$, and according to Cauchy–Schwarz inequality, we have

$$\|\nu_{\mathcal{S}_0^c}\|_1 \leq \|\nu_{\mathcal{S}_0}\|_1 \leq \sqrt{S} \|\nu_{\mathcal{S}_0}\|_2.$$

Thus, we have

$$\begin{aligned} \|\nu\|_1 &= \|\nu_{\mathcal{S}_0^c}\|_1 + \|\nu_{\mathcal{S}_0}\|_1 \\ &\leq 2\|\nu_{\mathcal{S}_0}\|_1 \\ &\leq 2\sqrt{S} \|\nu_{\mathcal{S}_0}\|_2 \\ &\leq 2\sqrt{S} \|\nu\|_2. \end{aligned} \quad (32)$$

For $\frac{1}{n}\|X^T X\nu\|_\infty$, we have

$$\begin{aligned}\frac{1}{n}\|X^T X\nu\|_\infty &= \max_{i=1}^p \left| \left\langle \frac{1}{\sqrt{n}} X_i, \frac{1}{\sqrt{n}} X\nu \right\rangle \right| \\ &\leq \max_{i=1}^p \left\| \frac{1}{\sqrt{n}} X_i \right\|_1 \left\| \frac{1}{\sqrt{n}} X\nu \right\|_\infty \\ &\leq \left\| \frac{1}{\sqrt{n}} X \right\|_1 \left\| \frac{1}{\sqrt{n}} X\nu \right\|_\infty.\end{aligned}\tag{33}$$

According to our Formulation (10), and the fact that both $\hat{\beta}$ and β^* are feasible to (10), we have

$$\left\| \frac{1}{\sqrt{n}} X\nu \right\|_\infty \leq 2\lambda_0.$$

According to the normalization on X and Cauchy-Schwarz inequality, we have

$$\left\| \frac{1}{\sqrt{n}} X \right\|_1 \leq O(\sqrt{n}).\tag{34}$$

Combing the upper-bounds above, we have

$$\begin{aligned}\gamma \|\nu\|_2^2 &\leq \frac{1}{n} \|X\nu\|_2^2 \leq \frac{1}{n} \|X^T X\nu\|_\infty \|\nu\|_1 \\ &\leq O(\lambda\sqrt{Sn}) \|\nu\|_2.\end{aligned}\tag{35}$$

Thus we obtain that

$$\|\nu\|_2 \leq O(\sqrt{S \log p}).\tag{36}$$

□

A.3 Proof of Theorem 3.5

Proof. Recall the way we bound $\left\| \frac{1}{\sqrt{n}} X \right\|_1$ in Inequality (34), if we in addition have that $\|X\|_1 = O(\sqrt{n})$, we have

$$\begin{aligned}\gamma \|\nu\|_2^2 &\leq \frac{1}{n} \|X\nu\|_2^2 \leq \frac{1}{n} \|X^T X\nu\|_\infty \|\nu\|_1 \\ &\leq O(\lambda\sqrt{S}) \|\nu\|_2.\end{aligned}\tag{37}$$

Thus we obtain that

$$\|\nu\|_2 \leq O\left(\sqrt{\frac{S \log p}{n}}\right).\tag{38}$$

□

A.4 Proof of Theorem 3.6

Proof. The proof simply follows from the previous proofs:

$$\begin{aligned}
 \frac{1}{n} \|X\nu\|_2^2 &\leq \frac{1}{n} \|X^T X\nu\|_\infty \|\nu\|_1 \\
 &\leq O(\lambda\sqrt{S}) \|\nu\|_2 \\
 &= O\left(\frac{S \log p}{n}\right).
 \end{aligned} \tag{39}$$

□

A.5 Proof of Proposition 4.1

Proof. In order to make the proof more clear, we use the equivalently statistical form, i.e., Problem (6) for the proof.

We will start with the simplest case where $p = 1$. In this case, we have the objective as

$$\min \|Y - X\beta\|_\infty + \lambda|\beta|, \tag{40}$$

where $Y, X \in \mathbb{R}^n$, $\beta \in \mathbb{R}$. By the definition of the ℓ_∞ norm, we can further write the objective in (40) as

$$\begin{aligned}
 &\min \|Y - X\beta\|_\infty + \lambda|\beta| \\
 &= \min \max_{i=1, \dots, n} \{|Y_i - X_i\beta|\} + \lambda|\beta|.
 \end{aligned}$$

It can be directly seen that this is a convex problem. For the first part $\max_{i=1, \dots, n} \{|Y_i - X_i\beta|\}$, which is the maximum of a set of convex functions, the resulting function is still convex. The second part $\lambda|\beta|$ itself is convex. Finding the minimum to (40) is equivalent to find a stationary solution to it.

On the one hand, for any $\beta_0 > 0$, we have $\lambda|\beta| = \lambda\beta$. Let $i \in I = \{i : |Y_i - X_i\beta_0| = \|Y - X\beta_0\|_\infty\}$, we have in a small neighborhood of β_0 , $\|Y - X\beta\|_\infty = |Y_i - X_i\beta|$. Thus in this small neighborhood $[\beta_-, \beta_+]$, where $\beta_-, \beta_+ \geq 0$, the objective is simply

$$f(\beta) = |Y_i - X_i\beta| + \lambda\beta.$$

By first order condition of the above function $f(\beta)$, we have

$$\partial f(\beta) = \lambda - \text{sign}(Y_i - X_i\beta)X_i.$$

According to the assumption on λ , the function $f(\beta)$ is increasing in the small neighborhood $[\beta_-, \beta_+]$.

On the other hand, for any $\beta_0 < 0$, we have $\lambda|\beta| = -\lambda\beta$. Let $i \in I = \{i : |Y_i - X_i\beta_0| = \|Y - X\beta_0\|_\infty\}$, we have in a small neighborhood of β_0 , $\|Y - X\beta\|_\infty = |Y_i - X_i\beta|$. Thus in this small neighborhood $[\beta_-, \beta_+]$, where $\beta_-, \beta_+ \leq 0$, the objective is simply

$$f(\beta) = |Y_i - X_i\beta| - \lambda\beta.$$

By first order condition of the above function $f(\beta)$, we have

$$\partial f(\beta) = -\lambda - \text{sign}(Y_i - X_i\beta)X_i.$$

According to the assumption on λ , the function $f(\beta)$ is decreasing in the small neighborhood $[\beta_-, \beta_+]$.

Using the two observations above, we can obtain that $\beta = 0$ is the unique global minimum for the objective in (40).

Now, we will generalize the above proof in the 1D case to high-dimensional case, where we assume $\beta \in \mathbb{R}^p$, $p > 1$.

For any $\beta_0 \in \mathbb{R}^p$, let $J = \{j : \beta_{0j} \neq 0\}$. Define the neighborhood of β_0 as $\mathcal{N} = \{\beta \in \mathbb{R}^p : \beta_j = 0, \text{ for } j \notin J; \beta_j \in [\beta_{0j} - \epsilon, \beta_{0j} + \epsilon], \text{ for } j \in J\}$, where $\epsilon > 0$ is chosen such that $(\beta_{0j} - \epsilon)(\beta_{0j} + \epsilon) > 0$ for all $j \in J$.

Let $i \in I = \{i : |Y_i - X_i\beta_0| = \|Y - X\beta_0\|_\infty\}$, we have that in the small neighborhood \mathcal{N} of β_0 , $\|Y - X\beta\|_\infty = |Y_i - X_i\beta|$. Thus within this small neighborhood \mathcal{N} , the objective is simply

$$f(\beta) = |Y_i - X_i\beta| + \lambda\|\beta\|_1.$$

By first order condition of the above function $f(\beta)$, we have

$$\partial f(\beta) = \lambda \text{sign}(\beta) - \text{sign}(Y_i - X_i\beta)X_i,$$

where $\text{sign}(\beta) \in \mathbb{R}^p$ is the indicator vector of the signs for β . According to the assumption on λ , $(\partial f(\beta))_k > 0$ for $k \in \mathbb{K}_1 = \{k : \beta_{0k} > 0\}$; $(\partial f(\beta))_k < 0$ for $k \in \mathbb{J}/\mathbb{K}_1$. Thus we conclude that $\beta = 0$, is the unique global minimum for the objective in (10), which is $[F'_1, F'_2]' = 0$ in the Problem (4). \square

B Algorithm 2

1: Input: $\rho, \lambda, e_3, e_4, n, M, B, U_1, \psi_1, U_2, \psi_2, A, b$
2: Initialize z^0, u^0
3: *loop*:
4: **for** $k = 1 : K$ **do**
5: $u_1^k = \{u_i^k\}_{i=1, \dots, 2n}, u_3^k = \{u_i^k\}_{i=1+2n, \dots, 2n+M},$
6: $z_1^k = \{z_i^k\}_{i=1, \dots, 2n}, z_3^k = \{z_i^k\}_{i=1+2n, \dots, 2n+M},$
7: $y_1^{k+1} = z_1^k - u_1^k - P_{B_{\|\cdot\|_1[0,1]}}(z_1^k - u_1^k),$
8: $y_3^{k+1} = z_3^k - u_3^k,$
9: $y^{k+1} = [(y_1^{k+1})', (y_3^{k+1})']',$
10: $z^{k+1} = y^{k+1} + u^k - E'(EE')^{-1}(Ez - b),$
11: $u^{k+1} = u^k + y^{k+1} - z^{k+1},$
12: $r^{k+1} = y^{k+1} - z^{k+1},$
13: $s^k = \rho(z^{k+1} - z^k),$
14: $e_1 = \sqrt{m}e_3 + e_4 \max\{\|z^{k+1}\|_2, \|y^{k+1}\|_2\},$
15: $e_2 = \sqrt{m}e_3 + e_4 \|\rho u^{k+1}\|_2,$
16: **if** $\|r^{k+1}\|_2 \leq e_1$ and $\|s^{k+1}\|_2 \leq e_2$ **then**
17: **return** y_3
18: **end if**
19: **end for**
Algorithm 2: The ADMM for estimating $y_3 = (y_2)_S = [(F_1)_{S_1}; (F_2)_{S_2}]$