

Frequency Domain Quasi Maximum Likelihood Identification of Low Order Aeroservoelastic Models from Flight-Test Data

Jeffrey Ouellette¹, Chris J. Miller², Matthew J. Boucher³

NASA Armstrong Flight Research Center, Edwards, California, 93523, U.S.A.

In this paper a quasi maximum likelihood method for estimating a low order model of a flexible vehicle has been developed and demonstrated. The quasi maximum likelihood method uses a large number of sensors to estimate the parameters and covariance in a method consistent with the maximum likelihood filter-error method. The cost function has been defined in a frequency domain to improve computational efficiency and to enable estimation of an unstable plant. The method has been demonstrated using flight data from the National Aeronautics and Space Administration X-56A Multi-Utility Technology Testbed flex wing tests, which include unstable flutter modes. The method was able to effectively estimate the frequency and damping of the dynamics of the aircraft and generate a transfer function that is useful for control system evaluations.

I. Nomenclature

$\mathbf{B}$	input SVD basis
$\mathbb{C}^{m \times n}$	set of $m \times n$ complex-valued matrices
$\mathbf{C}$	output SVD basis
EM	expectation-maximization
$\mathbf{H}_{loes}$	low order equivalent system transfer function
LOES	low order equivalent system
$\mathbf{H}_{\mathcal{L}}$	Hessian of the likelihood
$\mathbf{J}_{\mathcal{L}}$	Jacobian of the likelihood
MUTT	Multi-Utility Technology Testbed
$\mathbf{N}_j$	transfer function numerator coefficient matrix
OLS	ordinary least squares
$\hat{\mathbf{R}}$	input factor scores
$\mathbb{R}^{m \times n}$	set of $m \times n$ real-valued matrices
$\mathbf{S}_i$	reduced transfer function numerator coefficient matrix
$\mathbf{S}_u$	input SVD singular values
$\mathbf{S}_y$	output SVD singular values
SVD	singular value decomposition
$\mathbf{U}$	inputs in frequency domain
$U(0, \pi)$	uniform random variable from 0 to π
$\mathbf{v}_i$	outputs (sensor) noise
$\mathbf{V}_u$	normalized input basis vectors
$\mathbf{V}_y$	normalized output basis vectors
$\mathbf{w}_i$	inputs (process) noise vector
$\hat{\mathbf{X}}$	estimate of states

¹Aerospace Engineer, Controls and Dynamics Branch, P.O. Box 273/Mailstop 4840D, Edwards, California, 93523, AIAA Member.

²Former X-56 Chief Engineer, Controls and Dynamics Branch, P.O. Box 273/Mailstop 4840D, Edwards, California, 93523.

³Aerospace Engineer, Controls and Dynamics Branch, P.O. Box 273/Mailstop 4840D, Edwards, California, 93523.

$\mathbf{Y}$	outputs in frequency domain
$\hat{\mathbf{Y}}$	estimate of outputs
$\mathbf{Z}$	measured outputs in frequency domain

Subscripts

m_u	rank of inputs
m_y	rank of outputs
n	number of samples
p_u	number of inputs
p_y	number of outputs

Superscripts

H	Hermitian (complex conjugate) transpose
-----	---

Symbols

ζ_j	damping ratio of j -th mode
λ_{lm}	Levenberg-Marquardt parameter
σ_u	input standard deviation diagonal matrix
σ_y	output standard deviation diagonal matrix
ω_i	i -th frequency
$\boldsymbol{\theta}$	vector of model parameters
$\mathcal{V}$	matrix of innovations
Σ_ν	covariance of the innovations
Σ_v	outputs (sensor) noise covariance
Σ_w	inputs (process) noise covariance
ω_{n_j}	natural frequency of j -th mode

Conventions

$\dagger$	pseudo-inverse
$\in$	is an element of
tr	matrix trace

II. Introduction

As closed loop flight controls became more common, a method of low order equivalent systems (LOES) was developed to convert the high order dynamics of the combined control laws and vehicle plant into a lower order plant for comparison to classical handling quality metrics [1]. Morelli [2] demonstrated equation error and output error estimation methods in the frequency domain for generating these LOES models from flight data. These LOES models were limited to rigid body dynamics. Typically, they further restrict the model to a single mode, e.g., short-period or dutch roll.

Using modern design tools to create more capable and efficient aircraft has meant that structures are becoming lighter and more flexible. These increases in flexibility cause more interaction between the traditional flight dynamics and the structural dynamics of the vehicle. The traditional methods of assessing the dynamics of a vehicle separate the disciplines of flight dynamics and structural dynamics. Flight dynamics methods assume that the vehicle is a rigid body; however, with modern and more flexible aircraft it becomes impossible to separate these dynamics. For rigid body dynamics, the models were high order due to control law and actuator dynamics. For these flexible vehicle models, the structural dynamics and unsteady aerodynamics make the model high order. Miller et al. [3] extended the idea of generating LOES to these high order aeroelastic systems. A time domain output error approach was used to generate these low order systems. The lower order model simplified the structural dynamics and unsteady aerodynamics to a form that was easier to interpret and to use for assessing control law performance. These aeroelastic LOES models were instrumental in the National Aeronautics and Space Administration (NASA) envelope expansion and control law developments of the X-56A (Lockheed Martin) (Bethesda, Maryland) Multi-Utility Technology Testbed (MUTT). However, a number of problems were discovered as the methods were applied to speeds above the flutter speed.

The time domain methodology worked quite well at low airspeeds when the airframe was stable; however, Miller et al. [3] had numerous difficulties with the time domain method when the vehicle was unstable. The time domain

identification requires an implementation of the flight control laws and is biased by nonlinearities in the actuator responses. The dependence of the method on the control laws also limits the robustness of the numerical optimization and can prevent convergence to a solution. As the vehicle became significantly unstable, the time domain solution would not work so a frequency domain method was considered.

The maximum likelihood method often uses a Kalman filter to estimate the states and is therefore also called filter error. Relatively few system identification studies have made use of the maximum likelihood methods [4, 5] because they typically suffer from numerical problems. Most successful attempts to use maximum likelihood methods have used some approximation to achieve satisfactory results. For example, the output error method [6] assumes that there is no process noise, i.e., $\Sigma_w \triangleq \mathbf{0}$. Alternatively, Grauer and Morelli [7] assume a fixed value for the sensor noise, Σ_v , is known a priori. For a typical rigid vehicle, which has relatively few sensors, these methods are very effective. However, as the number of outputs and inputs are increased, the number of parameters being fit also increases. This “curse of dimensionality” means that adding additional sensors provides almost no benefit to the accuracy of the fit. For the maximum likelihood methods that already suffer from numerical problems, the curse of dimensionality makes the system identification even more challenging and significantly more computationally expensive. Typically, maximum likelihood system identification has used a cost function defined on the raw time domain data [4, 7]. The time domain does facilitate the implementation of a Kalman filter. For the present analysis, a quasi maximum likelihood method has been developed as an approximation that is consistent with the filter error method without requiring a Kalman filter. As a result, either a time domain or frequency domain methodology is feasible.

The present frequency domain quasi maximum likelihood method was developed to address the problems with the unstable systems and to incorporate additional sensor outputs. First, the general low order model is described. Second, a fitting approach for the low order model is developed. The large number of inputs and outputs are reduced to a lower rank representation. The low order representation is a quasi maximum likelihood method and is shown to be consistent with a classical maximum likelihood method. Third, the proposed method has been applied to flight-test data from the X-56A flex wing program. A number of metrics are discussed to show the effectiveness of the low rank representation of the input and output data. The frequency and damping data for these models are used to summarize the dynamics of the vehicle. Finally, the accuracy of the LOES are demonstrated against nonparametric estimates of the vehicle frequency response.

III. Model Definition

Consider the general form of a linear model with inputs, $\mathbf{U}$, and measured outputs, $\mathbf{Z}$ in the frequency domain, Eqs. (1a) and (1b). For a specific frequency in the frequency domain, the system response is a linear relationship between the inputs and the outputs. The linearity of the input-output relationship is true even for some nonlinear, e.g., nonrational, transfer functions. The optimization is still nonlinear in the parameters. In the frequency domain, the solution is always bounded regardless of stability, so no implementation of the control law is needed.

The general linear model of Eqs. (1a) and (1b) also includes process noise, Eq. (1c), and sensor noise, Eq. (1d). The noise is defined as a uniformly distributed phase angle and a constant fixed magnitude. This noise definition in the frequency domain is the equivalent of Gaussian white noise in the time domain.

$$\mathbf{Y}_{(p_y \times 1)}(\omega_i) = \mathbf{H}_{loes}(\omega_i | \boldsymbol{\vartheta})_{(p_y \times p_u)} \left(\mathbf{U}_{(p_u \times 1)}(\omega_i) + \sum_{(p_u \times p_u)}^{\frac{1}{2}} e^{i\mathbf{w}_i} \right) \quad (1a)$$

$$\mathbf{Z}_{(p_y \times 1)}(\omega_i) = \mathbf{Y}_{(p_y \times 1)}(\omega_i) + \sum_{(p_y \times p_y)}^{\frac{1}{2}} e^{i\mathbf{v}_i} \quad (1b)$$

$$\mathbf{w}_i_{(p_u \times 1)} \sim U(0, \pi) \quad (1c)$$

$$\mathbf{v}_i_{(p_y \times 1)} \sim U(0, \pi) \quad (1d)$$

The classical form of a LOES used in rigid body aircraft handling qualities [1, 2] can be generalized to the form of Eq. (2). The dynamics are separated into longitudinal and lateral axes. For a rigid vehicle, the transfer function is 2nd or 4th order, $m = 1$ or 2 . For the longitudinal axis $m = 2$, the two modes are the phugoid (airspeed and altitude) and

short-period (pitch) modes. To consider an aeroelastic vehicle, the value of m will be larger.

$$\mathbf{H}_{loes}(\omega_i|\boldsymbol{\vartheta})_{(p_y \times p_u)} \triangleq \frac{\sum_{j=0}^{2m} \mathbf{N}_j (i\omega_i)^j}{\prod_{j=1}^m (2i\zeta_j \omega_{n_j} \omega_i + \omega_{n_j}^2 - \omega_i^2)} \quad (2)$$

An alternative expression of the transfer function, Eq. (2), is shown in Eq. (3).

$$\mathbf{H}_{loes}(s|\boldsymbol{\vartheta})_{(p_y \times p_u)} = \frac{\mathbf{C} \left(\sum_{j=0}^{2m} \mathbf{S}_i (i\omega_i)^j \right) \mathbf{B}}{\prod_{j=1}^m (2i\zeta_j \omega_{n_j} \omega_i + \omega_{n_j}^2 - \omega_i^2)} \quad (3)$$

The alternative formulation, Eq. (4), is equivalent if the rank of $\mathbf{S}_i$ is equal to the rank of the smaller matrix $\mathbf{U}$ or $\mathbf{Y}$.

$$\mathbf{N}_i_{(p_y \times p_u)} \triangleq \frac{\mathbf{C} \quad \mathbf{S}_i \quad \mathbf{B}}{(p_y \times m)(m \times m)(m \times p_u)} \quad (4)$$

Where the input and output bases, $\mathbf{B}$ and $\mathbf{C}$, are defined to be orthogonal, Eqs. (5a) and (5b).

$$\mathbf{C}^T \mathbf{C} \quad \text{is diagonal} \quad (5a)$$

$$\mathbf{B} \mathbf{B}^T \quad \text{is diagonal} \quad (5b)$$

The orthogonality is not necessary, but ensures the uniqueness of the basis and improves the numerical properties of the fitting process. The alternative formulation of Eq. (3) allows one to reduce the number of variables being fit. Because there is only a single set of output bases, $\mathbf{B}$ and $\mathbf{C}$, the total number of terms in the matrices, $\mathbf{C}$, $\mathbf{S}_i$, and $\mathbf{B}$, can be less than the number of terms in all the $\mathbf{N}_j$ matrices.

IV. Fitting Methodology

The innovation, Eq. (6), is defined as the error between the model and the measured outputs.

$$\boldsymbol{\nu}_{(p_y \times n)} \triangleq \mathbf{Z}_{(p_y \times n)} - \hat{\mathbf{Y}}_{(p_y \times n)} \quad (6)$$

Where the outputs, $\hat{\mathbf{Y}}$, are estimated from the state equation without the process noise, Eq. (7).

$$\hat{\mathbf{Y}}_{(p_y \times 1)}(\omega_i|\boldsymbol{\vartheta}) = \mathbf{H}_{loes}(\omega_i|\boldsymbol{\vartheta})_{(p_y \times p_u)} \mathbf{U}_{(p_u \times 1)}(\omega_i) \quad (7)$$

The likelihood that the measured data came from the model defined by the parameters, $\boldsymbol{\vartheta}$, is given by Eq. (8) [8].

$$\mathcal{L}(\mathbf{Z}|\boldsymbol{\vartheta}) = \frac{1}{(2\pi)^{\frac{np_y}{2}} |\boldsymbol{\Sigma}_\nu|^{\frac{n}{2}}} \exp \left[-\frac{1}{2} \text{tr} \left(\boldsymbol{\nu}^H \boldsymbol{\Sigma}_\nu^{-1} \boldsymbol{\nu} \right) \right] \quad (8)$$

The negative log likelihood, Eq. (9), is used to convert to a minimization and simplify the equations.

$$-\ln \mathcal{L}(\mathbf{Z}|\boldsymbol{\vartheta}) = \frac{1}{2} \left[\text{tr} \left(\boldsymbol{\Sigma}_\nu^{-1} \boldsymbol{\nu} \boldsymbol{\nu}^H \right) + n \ln |\boldsymbol{\Sigma}_\nu| \right] + \text{constant} \quad (9)$$

The problem of finding the parameters, $\boldsymbol{\vartheta}$, that maximize likelihood, Eq. (8), is the general maximum likelihood system identification problem [4, 5].

To avoid the curse of dimensionality, the problem is separated into two parts. First, a set of basis vectors, $\mathbf{V}_u$ and $\mathbf{V}_y$, are generated from the data without any knowledge of the system dynamics. Rather than the system dynamics, the determination of these basis vectors uses the spatially distributed sensors to look at patterns at a single time and is achievable because of the large number of outputs. Instead of becoming a computational hindrance, the large number of outputs becomes a benefit. These basis vectors can be calculated using robust and fast numerical methods. Second, is to estimate the system dynamics parameters, ζ_i , ω_{n_i} , and $\mathbf{S}_i$. The application of the basis vectors, Eq. (3), in the previous step reduces the size of the $\mathbf{S}_i$ matrix and the number of parameters fit in the second step. Doz et al. [9] called this a quasi maximum likelihood method and showed that as the length of time and the number of outputs are increased the quasi maximum likelihood method is consistent with a true maximum likelihood method. This two-stage approach is similar to the factor score regression method [10, 11]. The primary difference is that factor score regression has a linear algebraic equation rather than a differential equation as the underlying latent process.

A. Principal Factor Analysis of Complex Data

In principal factor analysis [8], the input and output variables, $\mathbf{U}$ and $\mathbf{Z}$, are assumed to be a linear combination of m_u or m_y latent variables or modes. These outputs are standardized by the diagonal standard deviation matrices, σ_u and σ_y . These mode shapes, i.e., basis vectors, can be calculated from the matrix decomposition of Eqs. (10a) and (10b).

$$\begin{matrix} \mathbf{U} \\ (p_u \times n) \end{matrix} \triangleq \begin{matrix} \sigma_u & \mathbf{V}_u & \mathbf{S}_u & \hat{\mathbf{R}} \\ (p_u \times p_u) & (p_u \times m_u) & (m_u \times m_u) & (m_u \times n) \end{matrix} \quad (10a)$$

$$\begin{matrix} \mathbf{Y} \\ (p_y \times n) \end{matrix} \triangleq \begin{matrix} \sigma_y & \mathbf{V}_y & \mathbf{S}_y & \hat{\mathbf{X}} \\ (p_y \times p_y) & (p_y \times m_y) & (m_y \times m_y) & (m_y \times n) \end{matrix} \quad (10b)$$

It is important to note that the principal factor analysis, Eqs. (10a) and (10b), is actually an estimation of the covariance or correlation matrix of the data [8, 12]. Because we are using the standardized signals, the decomposition in Eqs. (10a) and (10b) is equivalent to a decomposition of the correlation matrix. The correlation matrix is used to address the scaling differences between dissimilar sensors such as accelerometers and strain gages.

Adding additional samples, $n \rightarrow \infty$, will improve the estimate of the covariance matrix, but will not necessarily increase the accuracy of the estimation of the basis vectors, $\mathbf{B}$ and $\mathbf{C}$. The input basis, $\mathbf{B}$, is typically known from the design of the test, and its accuracy is well established. In contrast, the output basis, $\mathbf{C}$, is not well known. The factor analysis is averaging across all the sensors in a spatially-distributed sense and is a consistent estimator of the output basis; therefore, as the number of outputs is increased, $p_y \rightarrow \infty$, the output matrix, $\mathbf{C}$, will converge to the exact solution - even for a small number of samples. Furthermore, the estimate of the states, $\hat{\mathbf{X}}$, will also converge toward the exact solution. For a classical stiff aircraft, there may only be one or two outputs; therefore, the use of the factor analysis is of little merit. However, for an aircraft instrumented to sense the structure, there could be dozens or perhaps hundreds of sensors. From the present work, the availability of $\sim 10+$ outputs is sufficient to warrant the additional effort of the factor estimation.

In a case where you have real-valued data, the decomposition of Eqs. (10a) and (10b) can be calculated by a singular value decomposition. However, when the input and output variables are complex-valued, such as with frequency domain data, additional requirements on the matrices make the calculation more difficult. For complex data, the normalized basis vectors or mode shapes, $\mathbf{V}_u$ and $\mathbf{V}_y$, are constrained to be real-valued, Eqs. (11a) and (11b). The constraint ensures that the transfer function coefficients of Eq. (3) will remain real-valued. The diagonal matrices of singular values, $\mathbf{S}_u$ and $\mathbf{S}_y$, are a scaling factor that should change the magnitude and not the phase and therefore must be real-valued and diagonal, Eqs. (11c) and (11d). The phase should be captured by the input factor scores, $\hat{\mathbf{R}}$, and the states, $\hat{\mathbf{X}}$, and therefore they are the only matrices that should have complex values, Eqs. (11e) and (11f).

$$\mathbf{V}_u \in \mathbb{R}^{p_u \times m_u} \quad (\text{normalized}) \quad (11a)$$

$$\mathbf{V}_y \in \mathbb{R}^{p_y \times m_y} \quad (\text{orthonormal}) \quad (11b)$$

$$\mathbf{S}_u \in \mathbb{R}^{m_u \times m_u} \quad (\text{diagonal}) \quad (11c)$$

$$\mathbf{S}_y \in \mathbb{R}^{m_y \times m_y} \quad (\text{diagonal}) \quad (11d)$$

$$\hat{\mathbf{R}} \in \mathbb{C}^{m_u \times n} \quad (\text{normalized}) \quad (11e)$$

$$\hat{\mathbf{X}} \in \mathbb{C}^{m_y \times n} \quad (\text{normalized}) \quad (11f)$$

For the inputs, the normalized basis is known from the design of the maneuver. For the present analysis, it is then sufficient to assume that the matrix of singular values is identity. The input scaling matrix is no longer truly singular values, but the terminology is retained. The input factor scores are then determined from Eq. (10a) using an ordinary least squares (OLS), or Bartlett method [8, 10, 13], Eq. (12).

$$\begin{matrix} \hat{\mathbf{R}} \\ (m_u \times n) \end{matrix} \triangleq \begin{matrix} \mathbf{S}_u^{-1} & \mathbf{V}_u^\dagger & \sigma_u^{-1} & \mathbf{U} \\ (m_u \times m_u) & (m_u \times p_u) & (p_u \times p_u) & (p_u \times n) \end{matrix} \quad (12)$$

For the output signals, $\mathbf{Z}$, the generalized singular value decomposition problem is solved by separating the real and imaginary parts of the signals. The real and imaginary parts are stacked into an augmented matrix in Eq. (13).

$$\begin{bmatrix} \text{real}(\mathbf{Z}) & \text{imag}(\mathbf{Z}) \end{bmatrix} \quad (p_y \times 2n) \quad (13)$$

The singular value decomposition of Eq. (13) gives the value of the normalized output basis and singular values. Using the Bartlett method for the states would cause our final model to be biased due to the sensor noise [10, 11]. Instead, the states are determined from Eq. (10b) using a regression, or Thomson method [8, 10, 13], Eq. (14c). Using the regression method for the state estimation is known as a bias reduction method. The Thomson method begins by constructing an estimate of the state correlation matrix, $\hat{\mathbf{P}}$, in Eq. (14a). From the state correlation, the sensor noise, $\hat{\Sigma}_v$, is estimated by assuming that the noise is uncorrelated between sensors, Eq. (14b). The estimate of the states is then defined by Eq. (14c).

$$\underset{(m_y \times m_y)}{\hat{\mathbf{P}}} = \underset{(m_y \times p_y)}{\mathbf{V}_y^T} \underset{(p_y \times p_y)}{\boldsymbol{\sigma}_y^{-1}} \underset{(p_y \times p_y)}{\mathbf{Z}\mathbf{Z}^T} \underset{(p_y \times p_y)}{\boldsymbol{\sigma}_y^{-1}} \underset{(p_y \times m_y)}{\mathbf{V}_y} \quad (14a)$$

$$\underset{(p_y \times p_y)}{\hat{\Sigma}_v} = \text{diag} \left(\underset{(p_y \times p_y)}{\mathbf{Z}\mathbf{Z}^H} - \underset{(p_y \times p_y)}{\mathbf{V}_y \mathbf{S}_y \hat{\mathbf{P}} \mathbf{S}_y^T \mathbf{V}_y^T} \right) \quad (14b)$$

$$\underset{(m_y \times n)}{\hat{\mathbf{X}}} \triangleq \left(\underset{(m_y \times m_y)}{\mathbf{I} + \hat{\mathbf{P}} \mathbf{S}_y \mathbf{V}_y^T \hat{\Sigma}_v^{-1} \mathbf{V}_y \mathbf{S}_y} \right)^{-1} \underset{(m_y \times p_y)}{\mathbf{S}_y \mathbf{V}_y^T \hat{\Sigma}_v^{-1}} \underset{(p_y \times n)}{\mathbf{Z}} \quad (14c)$$

Similar to a Kalman filter, the regression method is a maximum likelihood estimate of the states; however, the regression method is neglecting the dynamics of the system. This method is equivalent to a Kalman filter where the system dynamics become fast, i.e., the poles tend toward negative infinity. The presence of sensor noise causes estimates of the states and the parameter values to be biased away from zero.

The basis vectors, Eqs. (15a) and (15b), are also defined from the normalized basis vectors.

$$\underset{(m_u \times p_u)}{\mathbf{B}} \triangleq \underset{(m_u \times m_u)}{\mathbf{S}_u} \underset{(m_u \times p_u)}{\mathbf{V}_u^T} \underset{(p_u \times p_u)}{\boldsymbol{\sigma}_u} \quad (15a)$$

$$\underset{(p_y \times m_y)}{\mathbf{C}} \triangleq \underset{(p_y \times p_y)}{\boldsymbol{\sigma}_y} \underset{(p_y \times m_y)}{\mathbf{V}_y} \underset{(m_y \times m_y)}{\mathbf{S}_y} \quad (15b)$$

When the outputs are complex-valued, the present approach will maintain the desired requirements for the basis vectors. However, the states will no longer be orthogonal, $\hat{\mathbf{P}}$ is not identity, for complex data. Another note of the principal factor analysis is that because the factor analysis does not require knowledge of the system dynamics, it is only necessary to perform the calculations once.

B. Quasi-Filter Error in the Frequency Domain

The parameters are determined using an expectation-maximization (EM) algorithm [14, 15]. The EM algorithm is a two-step algorithm. The first step updates the expectation of the innovation noise, Σ_ν . The second step is the minimization of the negative log likelihood function, $-\ln \mathcal{L}(\mathbf{Z}|\boldsymbol{\vartheta})$. The minimization is equivalent to maximizing the likelihood. The two steps are repeated until the negative log likelihood function has converged. Generally, convergence is robust for the problems being examined here [15]; however, convergence can be arbitrarily slow and can converge to a local minimum. For the present work, convergence was typically less than 10 iterations of the EM algorithm.

For the expectation step, the expected value can be evaluated analytically. The condition in Eq. (16a) can be solved to give the expectation of the noise in Eq. (16b).

$$\frac{d}{d\Sigma_\nu} \ln \mathcal{L}(\mathbf{Z}|\boldsymbol{\vartheta}) = \left[\Sigma_\nu^{-1} \mathbf{v} \mathbf{v}^H \Sigma_\nu^{-1} - n \Sigma_\nu^{-1} \right] = 0 \quad (16a)$$

$$\implies \Sigma_\nu = \frac{1}{n} \mathbf{v} \mathbf{v}^H \quad (16b)$$

The innovations in Eq. (6) are approximated in Eq. (17) by the expressions for $\mathbf{Z}$ and $\mathbf{Y}$ from Eqs. (1b) and (10b).

$$\underset{(p_y \times n)}{\mathbf{v}} \approx \left(\underset{(p_y \times p_y)}{\boldsymbol{\sigma}_y} \underset{(p_y \times m_y)}{\mathbf{V}_y} \underset{(m_y \times m_y)}{\mathbf{S}_y} \underset{(m_y \times n)}{\hat{\mathbf{X}}} + \underset{(p_y \times p_y)}{\Sigma_v^{\frac{1}{2}}} \underset{(p_y \times n)}{e^{i\mathbf{v}}} \right) - \underset{(p_y \times n)}{\hat{\mathbf{Y}}} \quad (17)$$

To express the covariance of the innovations, Eq. (16b), in terms of known variables, the approximation of the innovation, Eq. (17), is used to give Eq. (18).

$$\Sigma_\nu \approx \frac{1}{n} \left(\underset{(p_y \times p_y)}{\boldsymbol{\sigma}_y} \underset{(p_y \times m_y)}{\mathbf{V}_y} \underset{(m_y \times m_y)}{\mathbf{S}_y} \underset{(m_y \times n)}{\hat{\mathbf{X}}} - \underset{(p_y \times n)}{\hat{\mathbf{Y}}} \right) \left(\underset{(p_y \times p_y)}{\boldsymbol{\sigma}_y} \underset{(p_y \times m_y)}{\mathbf{V}_y} \underset{(m_y \times m_y)}{\mathbf{S}_y} \underset{(m_y \times n)}{\hat{\mathbf{X}}} - \underset{(p_y \times n)}{\hat{\mathbf{Y}}} \right)^H + \Sigma_v \quad (18)$$

The maximization of the likelihood function cannot be solved analytically. The gradient of Eq. (7), $\frac{d\hat{\mathbf{Y}}_i}{d\boldsymbol{\vartheta}}$, is determined by finite differences to calculate the Jacobian, $\mathbf{J}_{\mathcal{L}}$, of Eq. (9) in Eq. (19a). The Hessian, Eq. (19b), is approximated from the Jacobian. The approximation of the Hessian is neglecting higher order terms but is generally sufficiently accurate to improve the rate of convergence of the parameter estimates.

$$\mathbf{J}_{\mathcal{L}}(\boldsymbol{\vartheta}) = -\sum_i^n \left(\frac{d\hat{\mathbf{Y}}_i}{d\boldsymbol{\vartheta}}^H \boldsymbol{\Sigma}_{\nu}^{-1} \hat{\mathbf{Y}}_i \right) \quad (19a)$$

$$\mathbf{H}_{\mathcal{L}}(\boldsymbol{\vartheta}) \approx \sum_i^n \left[\frac{d\hat{\mathbf{Y}}_i}{d\boldsymbol{\vartheta}}^H (\boldsymbol{\Sigma}_{\nu}^{-1} + \lambda_{lm} \mathbf{I}) \frac{d\hat{\mathbf{Y}}_i}{d\boldsymbol{\vartheta}} \right] \quad (19b)$$

The λ_{lm} parameter is the Levenberg-Marquardt parameter to constrain the step size and improve the robustness of the algorithm [16]. As the $\lambda_{lm} \rightarrow 0$, the Levenberg-Marquardt method becomes the typical quasi-Newton method. As the parameter is increased, $\lambda_{lm} \rightarrow \infty$, the step size decreases and shifts toward the steepest descent solution. Decreasing the step size will help improve the robustness. By shifting the direction toward the steepest descent also helps to improve robustness and ensure that the step will decrease the cost. The penalty is an increase in the required number of iterations. The step size can be scaled based on the accuracy of the linearized step [16]. This adaptive step size will help speed up the solution but maintain robustness when needed.

For each iteration the maximization step, the value of the parameters are updated using Eq. (20).

$$\boldsymbol{\vartheta}_{k+1} = \boldsymbol{\vartheta}_k + \mathbf{H}_{\mathcal{L}}^{-1}(\boldsymbol{\vartheta}) \mathbf{J}_{\mathcal{L}}(\boldsymbol{\vartheta}) \quad (20)$$

For the EM algorithm, a single parameter update is sufficient if the likelihood is decreased, but additional iterations will help reduce the total calculations by reducing the number of updates to the expectation. The parameter update is repeated until the parameters or the value of the likelihood converge.

The uncertainty is estimated from the Cramér-Rao bounds, Eq. (21).

$$\text{cov}(\boldsymbol{\vartheta}) \geq \left[\sum_i^n \left(\frac{d\hat{\mathbf{Y}}_i}{d\boldsymbol{\vartheta}}^H \boldsymbol{\Sigma}_{\nu}^{-1} \frac{d\hat{\mathbf{Y}}_i}{d\boldsymbol{\vartheta}} \right) \right]^{-1} \quad (21)$$

In the time domain, the Cramér-Rao bounds can be biased small due to autocorrelation (i.e., colored) residuals. The residuals end up correlated due to modes excluded from the model and from turbulence. However, in the frequency domain, these residuals are decorrelated, but the noise is no longer uniform with frequency. Typically, for rigid body system identification, the bandwidth is sufficiently small that the noise is very close to uniform, and as a result, the Cramér-Rao bounds are minimally biased. Unfortunately, for the present aeroelastic system analysis, the bandwidth is significantly higher. As a result, the turbulence may not be as uniform. Furthermore, asymmetric and higher order symmetric modes can cause peaks in the noise. Nonuniform noise will bias the parameter uncertainty toward zero. For the current analysis, the bias in parameter uncertainty due to nonuniform noise, heteroscedasticity, has been neglected but is discussed further in the results.

V. Results

The X-56A MUTT shown in Fig. 1, was selected for applying the proposed quasi maximum likelihood fitting methodology. The X-56A MUTT was developed to improve tools for modeling and controlling vehicles with strong coupling between the flight dynamics and the structural dynamics [17]. To achieve these objectives, two reconfigurable centerbodies and a ground control station were developed. The centerbody was designed to house the critical vehicle systems and the wings could be swapped to achieve the objectives of the different flight phases. A total of one set of stiff wings (flutter outside the operational envelope) and three sets of flexible wings (flutter inside the operational envelope) were built. The flexible wings were designed so that the rigid body, flight dynamics, modes were a major participant in the flutter mechanism. The X-56A MUTT has a substantial amount of flight-test data at conditions above and below the flutter speed.



Fig. 1 The X-56A MUTT in flight (credit: NASA/Jim Ross).

The proposed frequency domain quasi maximum likelihood method was applied to a total of 44 separate flight conditions. Two test points, Table 1, from flights 16 and 34 were selected to show more details on the factors and frequency responses. These conditions were selected to represent the range of airspeed and fuel weight tested.

Table 1 Test case flight conditions.

Case	Equivalent airspeed, kn	Fuel weight, lb	Flight
1	80	62	16
2	110	23	34

A. Input and Output Selection

The X-56A planform with the sensors labeled is shown in Fig. 2. For the present analysis, all 10 z -axis accelerometers have been used. Also, a total of eight strain gages have been used. These strain gages were oriented to primarily sense the wing bending loads. These sensors were selected because they have high bandwidth and minimal signal processing which allows the sensor dynamics to be neglected. Some sensors such as the inertial navigation system have significant processing that introduces latency and dynamics that are independent of the aircraft dynamics. The combination of accelerometers and strain gages are effective at sensing the desired modes.

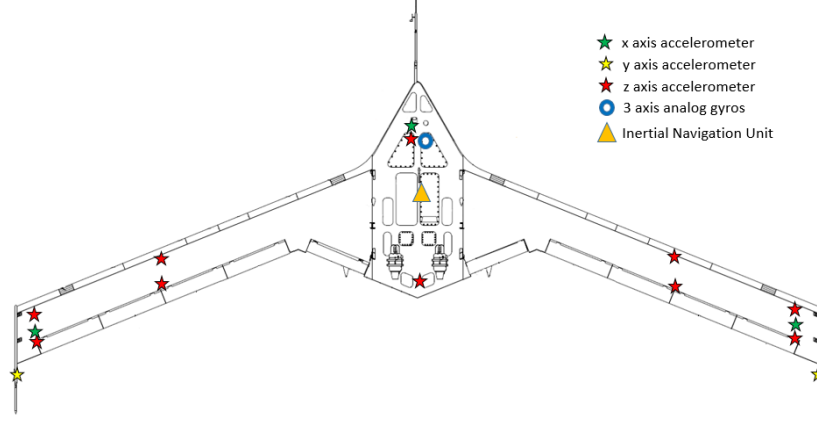


Fig. 2 The X-56A sensor and control surface locations.

B. Maneuver Description

The maneuvers used in the present analysis were all orthogonal phase-optimized multisines [5]. All 10 surfaces are excited in symmetric pairs by an orthogonal multisine input, basis in Eq. (22). The multisines are added directly to the command going from the flight computer to the actuator.

$$\mathbf{V}_u \triangleq \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \quad (22)$$

These maneuvers are a summation of sinusoids and are spaced to ensure that the inputs to each surface pair are orthogonal. The contributions of the control system, nonlinearities, and noise mean that the inputs are no longer completely orthogonal but still have low correlation. The symmetric maneuvers were intended to excite primarily the longitudinal flight dynamics and the symmetric structural modes. The frequency bandwidth was intended to excite the short-period, first symmetric wing bending, and first symmetric wing torsion modes. The magnitude of the maneuver was adjusted based on the flight condition to maximize the signal to noise without overloading the structure.

The current frequency domain system identification technique is also well suited to these multisine inputs. No windowing is used for the Fourier transform. For the multisine, the designed input frequencies fall directly at the lobes of the discrete Fourier transform. Any spectral leakage for a specific multisine frequency is at the minima of the other frequencies and is well attenuated.

C. Parameter Initial Estimate

The numerator coefficients are a linear relationship between the inputs and the outputs and as a result, the method is relatively insensitive to the initial value of $\mathbf{S}_i$. Therefore, a simple initial $\mathbf{S}_i = 0$ is used.

Convergence is more sensitive to the initial estimate for the frequency and damping. Various methods were tried to find a solution that would work best. A reasonable solution was found by trial and error, similar to Miller et al. [3]. However, there were a few flight conditions that had trouble converging. Furthermore, there were problems with the ordering of the modes changing frequently. It was difficult to determine when these problems would arise, so alternative methods of determining the initial estimate were explored. Noniterative system identification methods [18] were explored to generate an initial estimate of the frequency and damping; however, these methods are unable to recognize the stability of the system and always return a stable model. As a result, the final model would incorrectly converge to a stable local minimum when the vehicle was unstable. Furthermore, these methods often produced many “junk” poles that required additional processing to identify and remove. The best method for an initial estimate of the frequency, ω_{n_i} , and damping, ζ_i , was to select from the linear model predictions [19].

D. Factor Analysis Results

There are many different methods for determining the correct number of factors to use. For the present analysis, the selection of three factors was based on engineering judgment and the preflight analysis of the flutter mechanism. The order of the LOES models was selected to capture the short-period, symmetric wing bending, and symmetric wing torsion modes. The selection of these three modes can be validated by examining the contribution of each factor to the total outputs shown in Fig. 3. The variance of each factor is represented by the diagonal elements of $\mathbf{S}_y$ in Eq. (10b). These terms represent the contribution to the total correlation matrix, and therefore the average of all the factors is one.

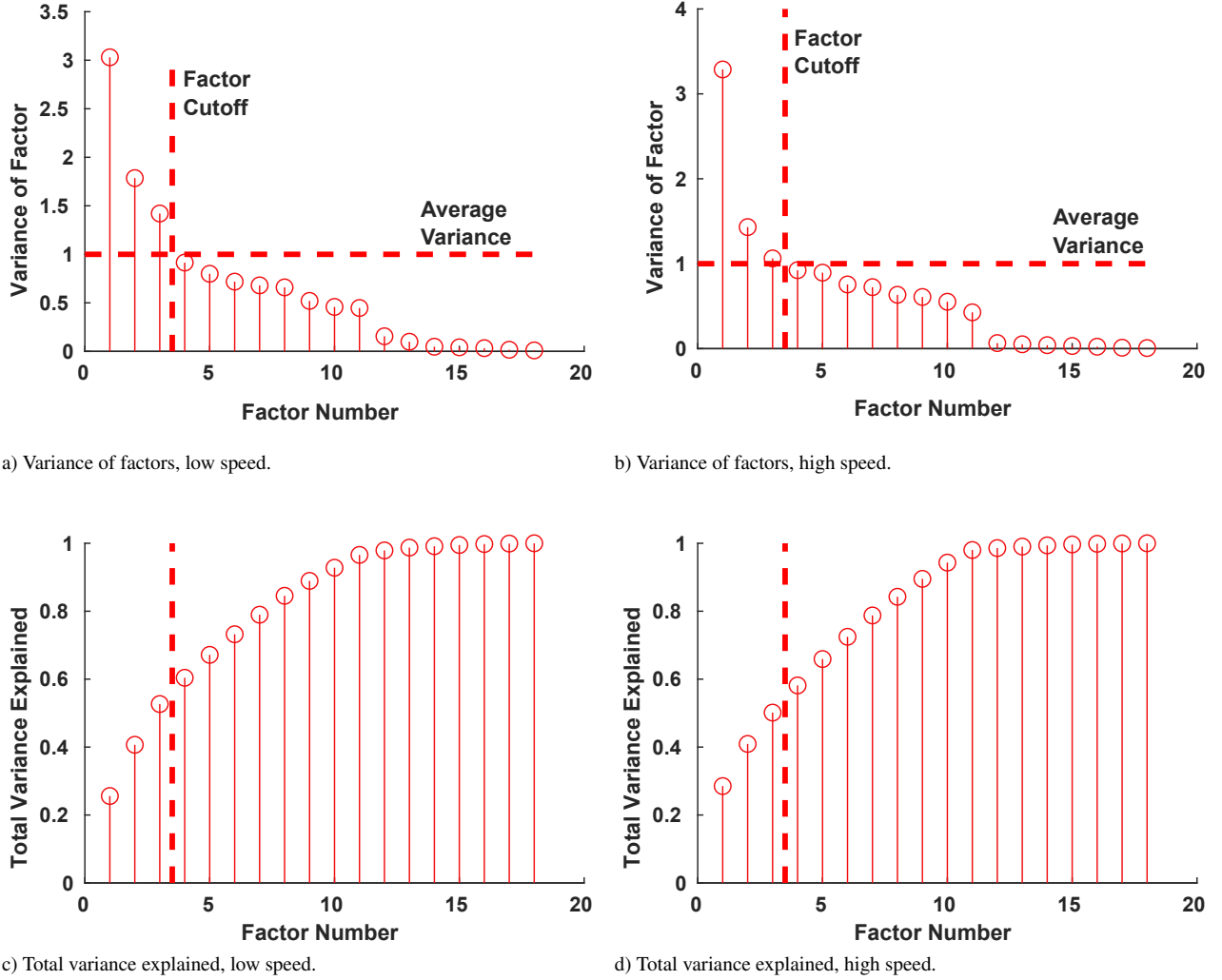


Fig. 3 Output factor scree plots.

One common rule for selecting the number of factors, Kaiser rule, is to select the factor with variance contribution greater than one [12]. Based on the Kaiser rule, the current assumption of three factors from engineering judgment is reasonable even though these three modes only explain roughly one-half of the variance. Another method is to look for a cutoff of the individual variances. For both these cases, there is a cutoff at 11 factors, which suggests it may be possible to create a model of higher order. However, the objective was to create a lower order model, which is more easily interpreted [1, 3]. Furthermore, as the order is increased, the number of parameters and the uncertainty, in turn, will increase.

The frequency content of each chosen state, $\hat{\mathbf{X}}$, is shown in Fig. 4 to demonstrate the adequacy of the number of factors. Due to the principal factor analysis, the factors are a function of patterns in the data across sensors and are not aeroelastic modes. As a result, individual factors can show peaks at multiple frequencies. However, these three factors

do show peaks primarily close to the frequency of the three aeroelastic modes of interest. If more factors were added to Fig. 4, then additional peaks, corresponding to higher order aeroelastic modes, would show up. To reduce the noise, it is desirable to keep the number of factors low to reduce the contributions of higher order structural dynamics.

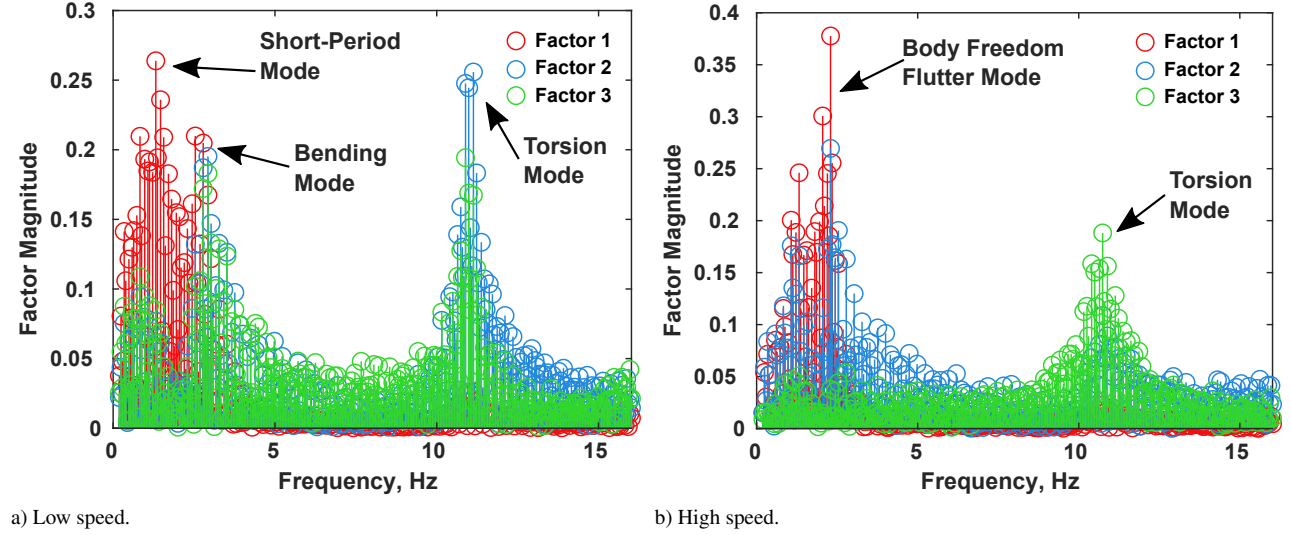


Fig. 4 Output frequency content for three chosen factors.

The final assessment of the factors is the normalized output basis, $\mathbf{V}_y$ of Eq. (10b), which is shown in Fig. 5. The normalized output basis shows the portion of each output explained by the factor, which can also be called the factor loading. For both flight conditions, a single factor is very effective at capturing nearly all of the strain gage content. Looking at the loadings, one can see that this first factor is primarily a symmetric wing bending mode. The other two factors are more difficult to interpret because they are a combination of the wing bending and torsion modes. The final column of the factor loadings, the communality, is the cumulative portion of the output explained by all of the factors being used. The communality is an estimate of the noise-to-signal ratio. A single factor was sufficient to explain nearly all of the strain gage results; therefore, the strain gage communality is nearly 100 percent. Most of the accelerometers are > 80 percent which indicates a good set of modes. The accelerometer at the nose and the aft midspan show the lowest communality but are still > 70 percent. These results indicate that more factors, and a higher order model may be needed to effectively capture the dynamics seen by these sensors. However, for the current analysis, the three factor model appears to be effective for both flight conditions shown here.

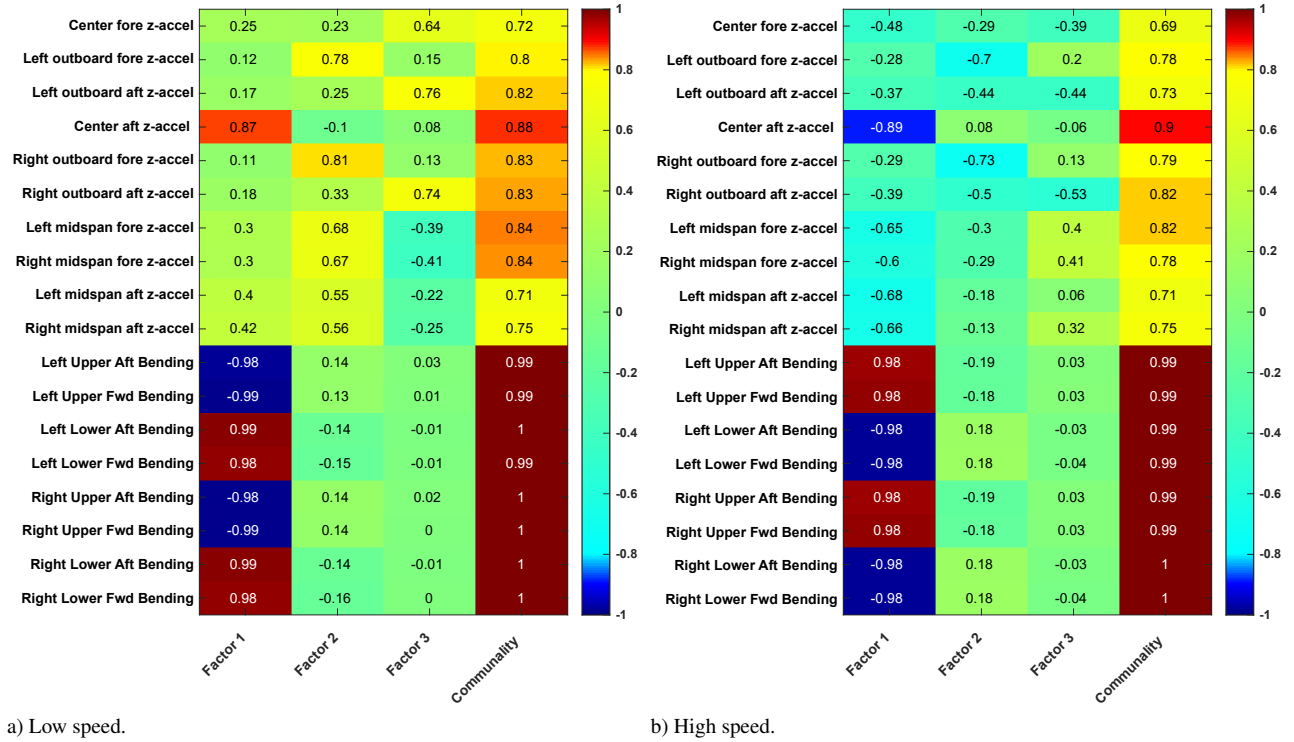


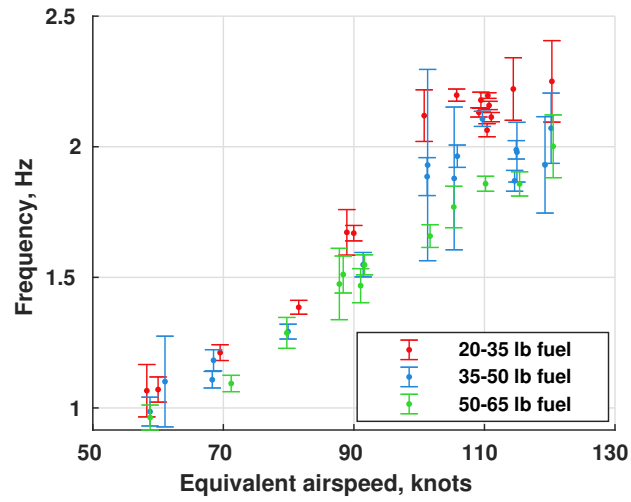
Fig. 5 Output factor loading.

The pattern of the normalized input basis is known from the experimental design, Eq. (22). However, nonlinearities in vehicle, asymmetries, and turbulence will cause anti-symmetric inputs that will cause the factor loadings to be < 1 . At the higher speeds, the input magnitude needs to be decreased, and there is more sensitivity to the turbulence. For both of these cases, the symmetric inputs are able to explain nearly all, > 98 percent, of the actuator commands.

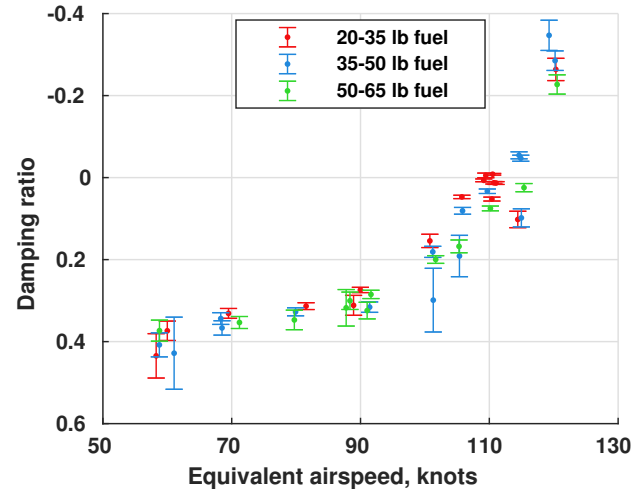
E. Frequency and Damping

The frequency and damping results for the three modes are summarized in Fig. 6. Unlike the factor analysis results in section V.D., “Factor Analysis Results,” these poles do represent the aeroelastic modes. A value of negative damping represents an unstable mode. The direction of the y -axis has been reversed for the damping to match the trends shown in typical flutter analyses. The fuel weights have been separated into a low (20-35 lb), mid (35-50 lb), and high (50-65 lb) range. The error bars are the 95 percent confidence bounds on the estimate. It is expected that the heteroscedasticity would cause these uncertainties to be biased small; however, they do appear to be reasonably sized compared to the dispersion of the estimates.

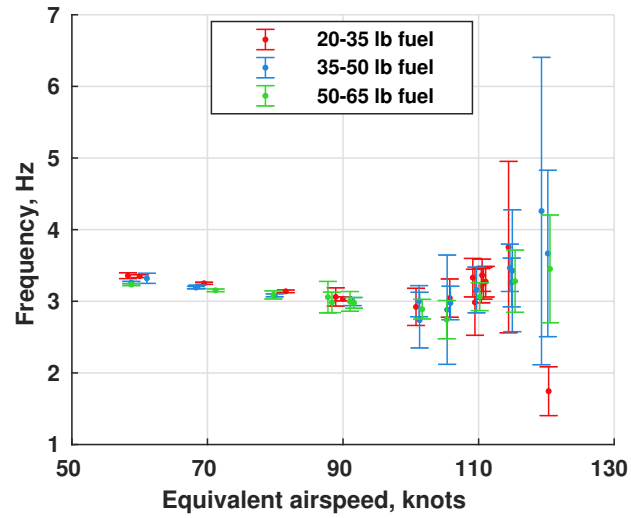
Overall, the present estimation method produced useful estimates of the frequency and damping of the modes. For the torsion modes, Figs. 6e and 6f, it appears that there are three cases at the maximum speed that did not converge correctly. Based on other analyses, there is another symmetric structural mode close to that frequency. The estimates may have converged to this higher frequency mode. For the other modes, there are some cases that may not have converged well but are less apparent than the torsional mode case. One can see that there is a significant increase in the uncertainty of the estimates as flight conditions get closer to the flutter speed. As the flutter mode becomes unstable, the sensitivity to turbulence is increased and the uncertainty is increased. Furthermore, nonlinearities in the actuator caused persistent small-magnitude limit cycle oscillation [20] at these speeds which further increased the uncertainty in the estimates.



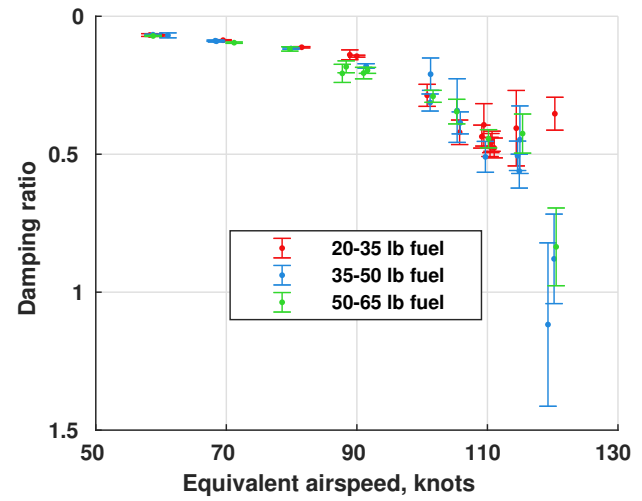
a) Body freedom frequency.



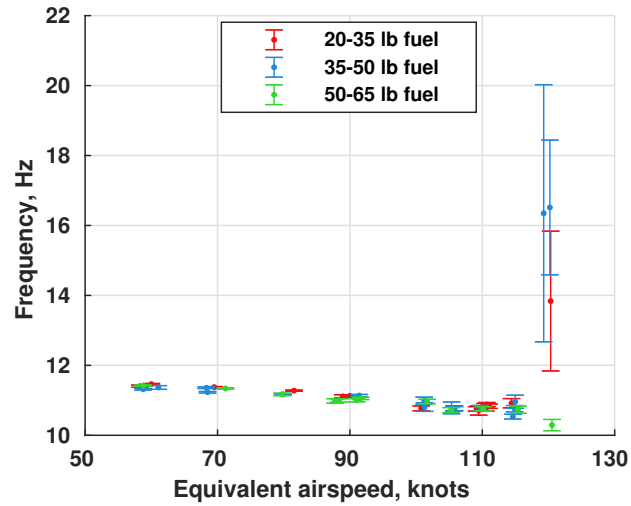
b) Body freedom damping.



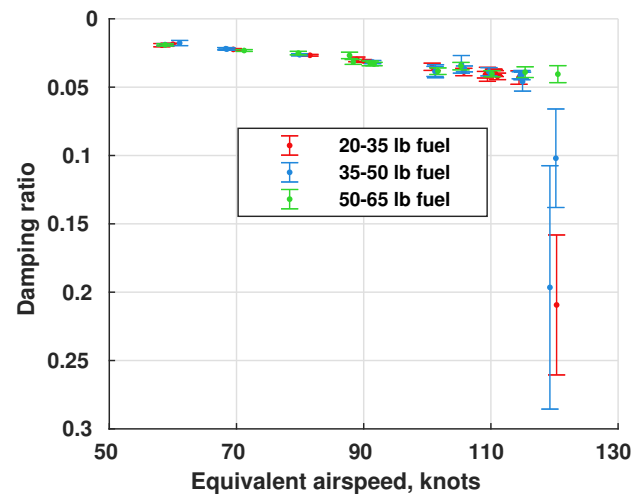
c) Symmetric wing bending frequency.



d) Symmetric wing bending damping.



e) Symmetric wing torsion frequency.



f) Symmetric wing torsion damping.

Fig. 6 Estimated mode frequency and damping as a function of airspeed.

The ability of the method to estimate the frequency and damping is useful for visualizing the trends with airspeed and fuel. The frequency of the body freedom flutter, Fig. 6a, is increasing with airspeed and decreasing with fuel weight. The damping of the body freedom flutter, Fig. 6b, shows that the instability occurs at approximately 110 to 120 knots. The effect of fuel on the damping of the body freedom flutter is quite small, but it does have a more noticeable change in the flutter speed. Other methodologies may show the flutter speed better [3] but were beyond the scope of the current work.

The effect of fuel weight on the frequency of the other modes, Figs. 6c and 6e, is much smaller and cannot be distinguished from the variance of the estimates. Looking at the symmetric wing bending mode, Figs. 6c and 6d, the increase in uncertainty at high speeds is even more pronounced. The mode is becoming overdamped and is very difficult to distinguish from the body freedom flutter mode which is very close in frequency. As a result, the uncertainty is very high. However, the modes are still statistically significant and are of value to leave in the model.

F. Frequency Responses

The frequency response for the center forward accelerometer and a strain gage are shown in Fig. 7 for the low-speed case from flight 16. The present low order model is compared to a nonparametric estimate of the frequency response using the method of Grauer and Boucher [21]. The input is symmetric excitation of the outboard-most flaps, which gives a good excitation of the longitudinal dynamics.

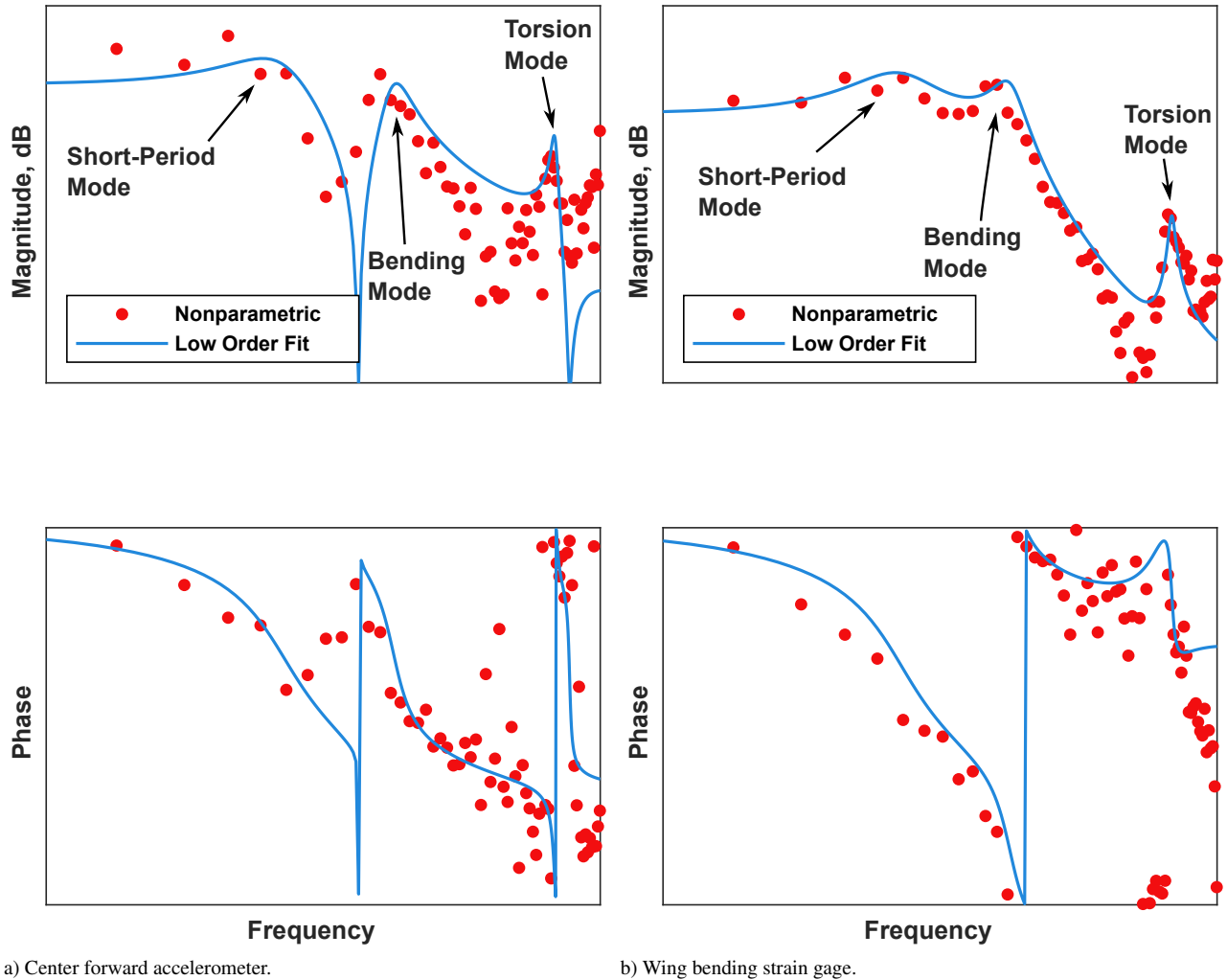


Fig. 7 Symmetric outboard surface excitations from flight 16.

At low speeds, the predictive model matches well to the flight-test data. The low order fit does show a difference in the phase of the accelerometer at the wing bending mode frequency. The low order model appears to have incorrectly identified the sign of the real part of a zero. The strain gage does show a good match. There is some difference in the matches between the wing bending and wing torsion modes. The poor fit is a limitation of the low order fit which does not include any unsteady aerodynamics or sensor dynamics. A higher order model would be required to better match the magnitude; however, these models do capture the primary dynamics and were sufficient for characterizing the control system design.

The frequency responses for the high-speed case, close to the flutter speed, in flight 34 are shown in Fig. 8. At the higher air speeds, there is more noise and the nonparametric model shows this noise. However, the phase does match well, and indicates a good fit for the present low order model. The limitation of the low order fit is more apparent at the higher speed where the nonparametric model shows a faster roll-off than can be achieved by a second-order mode, 20dB/decade.

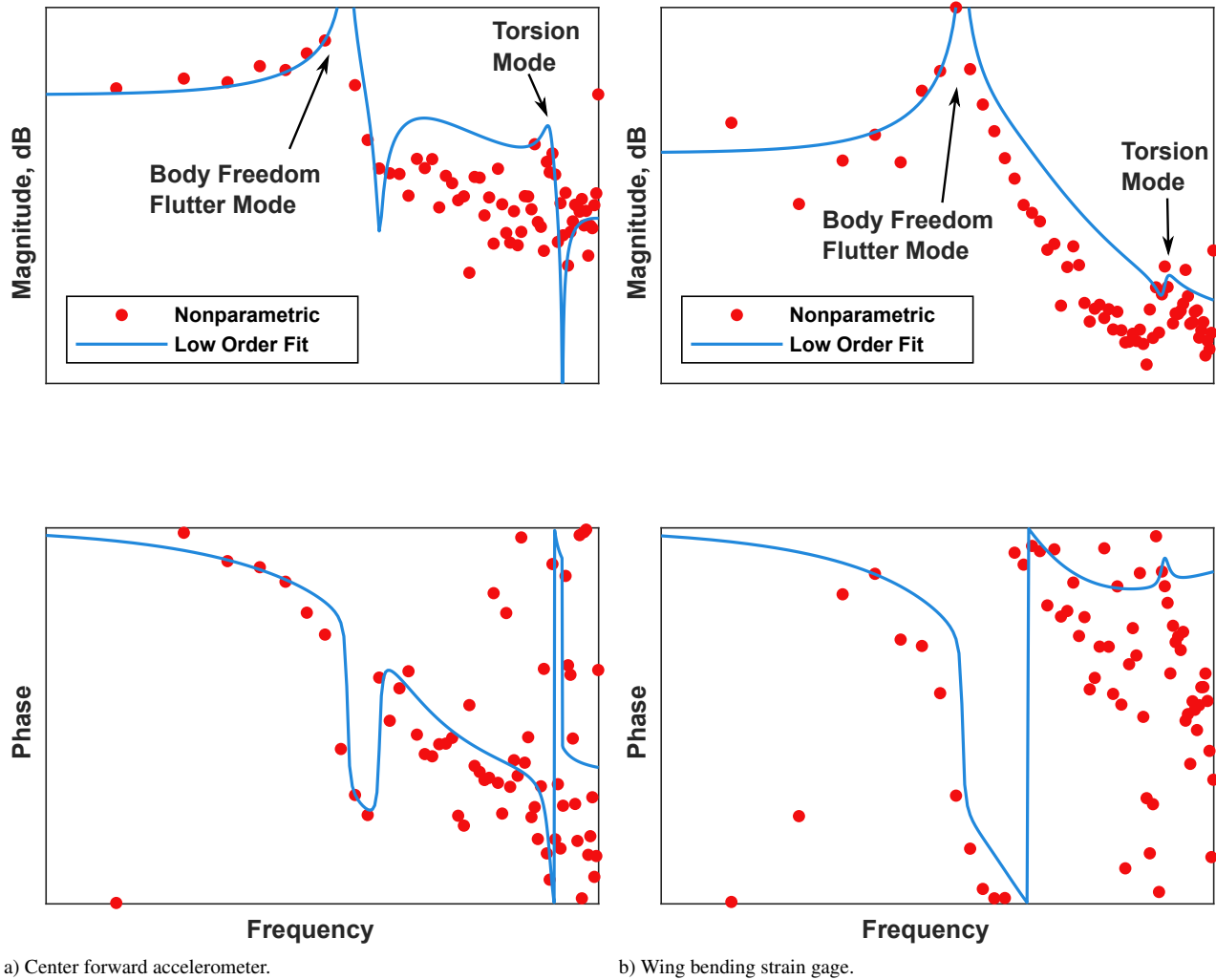


Fig. 8 Symmetric outboard surface excitations from flight 34.

VI. Conclusion

A quasi maximum likelihood method has been developed to estimate low order transfer functions to represent the coupled aeroelastic and flight dynamics from flight-test data. By neglecting the underlying equations of motion,

a factor model is used to estimate the states. The factor model is consistent with the more temperamental Kalman filter but can be more easily generated by instead exploiting a large number of outputs. The method has been used to generate low order models from the National Aeronautics and Space Administration X-56A Multi-Utility Technology Testbed flex wing flight tests for airspeeds above and below the flutter speed. The factor model was shown to be able to effectively reduce the number of outputs with a minimal loss of information, primarily sensor noise. The quasi maximum likelihood method was shown to effectively produce lower order models; however, the low order of the models does limit their accuracy. The models were useful for characterizing the dynamics through trends in the frequency and damping. Furthermore, the frequency responses matched well to high order nonparametric estimates of the frequency response functions. For a vehicle with a large number of sensors, the current approach provides a computationally efficient, reliable, and accurate method for estimating a low order unstable system.

Acknowledgments

This research was funded by the Air Force Research Laboratory, the NASA Advanced Air Transport Technology project, and the NASA Flight Demonstrations and Capabilities program. Thanks also go to all the members of the X-56A team whose efforts made the flight tests possible.

References

- [1] Mitchell, D. G., and Hoh, R. H., "Low-Order Approaches to High-Order Systems: Problems and Promises," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, 1982, pp. 482–489.
doi: 10.2514/3.56195
- [2] Morelli, E. A., "Identification of low order equivalent system models from flight test data," NASA/TM-2000-210117, 2000.
- [3] Miller, C. J., Schaefer, J., Boucher, M., Ouellette, J., and Howe, S., "X-56 Flight Test Approach for Envelope Expansion Past Open Loop Flutter Instability," STO-MP-AVT-211, 2022.
- [4] Maine, R. E., and Iliff, K. W., "Formulation and implementation of a practical algorithm for parameter estimation with process and measurement noise," 80-1603, 1980.
doi: 10.2514/6.1980-1603
- [5] Morelli, E. A., and Klein, V., *Aircraft System Identification: Theory and Practice*, Sunflyte Enterprises, Williamsburg, VA, 2016.
- [6] Maine, R. E., and Iliff, K. W., "AGARD Flight Test Techniques Series. Volume 3 on Identification of Dynamic Systems-Applications to Aircraft. Part 1: The Output Error Approach," AGARD-AG-300-VOL. 3 PART 1, AGARDograph No.300 Vol.3, Dec. 1986.
- [7] Grauer, J. A., and Morelli, E. A., "A new formulation of the filter-error method for aerodynamic parameter estimation in turbulence," AIAA 2015-2704, June 2015.
doi: 10.2514/6.2015-2704
- [8] Johnson, R. A., and Wichern, D. W., *Applied Multivariate Statistical Analysis*, Pearson Prentice Hall, 2007.
- [9] Doz, C., Giannone, D., and Reichlin, L., "A quasi maximum likelihood approach for large approximate dynamic factor models," European Central Bank, Working Paper Series 674, Sept. 2006.
- [10] Devlieger, I., Mayer, A., and Rosseel, Y., "Hypothesis testing using factor score regression: A comparison of four methods," *Educational and Psychological Measurement*, Vol. 76, No. 5, 2016, pp. 741–770.
- [11] Hoshino, T., and Bentler, P. M., "Bias in factor score regression and a simple solution," *UCLA Department of Statistics Papers*, Oct. 2011.
- [12] Lawley, D. N., and Maxwell, A. E., *Factor analysis as a statistical method*, Butterworth & Co. Ltd., 1963.
- [13] Thomson, G. H., "Methods of estimating mental factors," *Nature*, Vol. 141, No. 3562, 1938, pp. 246–246.
- [14] Dempster, A. P., Laird, N. M., and Rubin, D. B., "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 39, Issue. 1, 1977, pp. 1–22.
- [15] Wu, C. F. J., "On the convergence properties of the EM algorithm," *The Annals of statistics*, Vol. 11, No. 1, 1983, pp. 95–103.
- [16] Moré, J. J., "The Levenberg-Marquardt algorithm: implementation and theory," Springer, 1978, pp. 105–116.
- [17] Beranek, J., Nicolai, L., Buonanno, M., Burnett, E., Atkinson, C., Holm-Hansen, B., and Flick, P., "Conceptual Design of a Multi-Utility Aeroelastic Demonstrator," AIAA 2010-9350, 2010.
doi: 10.2514/6.2010-9350

- [18] Van Overschee, P., and De Moor, B., *Subspace identification for linear systems: Theory—Implementation—Applications*, Kluwer Academic Publishers, 1996.
- [19] Ouellette, J. A., and Valdez, F. D., “Generation and Calibration of Linear Models of Aircraft with Highly Coupled Aeroelastic and Flight Dynamics,” AIAA 2020-1016, 2020.
doi: 10.2514/10.2514/6.2020-1016
- [20] Pankonien, A. M., Suh, P. M., Schaefer, J. R., and Mitchell, R. M., “Deadbands Tell No Tails: X-56A Dynamic Actuation Requirements,” SMASIS2020-23159, Sept. 2020, V001T03A016.
doi: 10.1115/SMASIS2020-2427
- [21] Grauer, J. A., and Boucher, M. J., “Aircraft System Identification from Multisine Inputs and Frequency Responses,” AIAA 2020-0287, 2020.
doi: 10.2514/6.2020-0287