



# International Operational Modal Analysis Conference

20 - 23 May 2025 | Rennes, France

## Exploring Physically Meaningful Prior Distributions for OMA

*Brandon J. O'Connell*<sup>1</sup> and *Timothy J. Rogers*<sup>1</sup>

<sup>1</sup> Dynamics Research Group, School of Mechanical, Aerospace and Civil Engineering, University of Sheffield  
{b.j.oconnell@sheffield.ac.uk, tim.rogers@sheffield.ac.uk}

### ABSTRACT

Recent advances in Bayesian approaches to operational modal analysis (OMA) have motivated the need to explore a range of possible prior distributions over the parameters of the state space representation of linear time-invariant dynamic systems. Choosing prior structures that can guarantee physically meaningful solutions in the context of structural dynamics, a property not necessarily guaranteed in many state space model-based system identification methods (e.g. stochastic subspace identification), is of particular interest. In this work, a range of prior model definitions for state space models are explored, including those which embed and enforce existing knowledge of the physics, with the view that physically meaningful posterior estimates could be obtained. Four potential priors for the state matrix are shown, with samples of these priors visualised on the elements of the state matrix and the corresponding modal properties. The ability of these priors to satisfy the condition of providing physically meaningful estimates is then discussed. Finally, considerations or challenges when selecting such priors in the context of Bayesian inference are identified.

*Keywords: Bayesian, Priors, Physics-informed, Stochastic Subspace*

### 1. INTRODUCTION

State space models (SSMs) are regularly used in methods of system identification and have found particular success in output-only identification of structural dynamic systems and operational modal analysis (OMA) [1–4]. Of the available available methods, perhaps the most well known is the stochastic subspace identification (SSI) algorithm [5]; the now industry standard for performing time-series OMA. Many SSM-based approaches to system identification are general methods that can be used for solving the stochastic realisation problem for linear time-invariant (LTI) systems [5], in which structural dynamic systems are a subset. This property gives them broad applicability to any set of problems that can be described by an LTI SSM.

Despite their frequent usage, it is widely known that the construction of many SSM-based algorithms (e.g. SSI) do not principally restrict solutions to be physically meaningful in the context of structural dynamics, i.e. strictly positive damping ratios for a stable dynamic system. In practice non-physical estimates often appear during application to vibration data, particularly at higher model orders, with no mechanism to address this problem within the identification. Consequently, many estimates are often discarded after the fact as being spurious in nature, arising solely from the numerics of learning some set of possible SSM parameters.

With a rise in the number of probabilistic and Bayesian approaches to OMA for the purpose of uncertainty quantification, particularly those using SSMs [6–11], there is potential to limit the set of possible solutions to those that only obey the physics; through the inclusion of more informative prior structures.

In this brief study, the theory surrounding LTI SSMs is revisited and the theoretical definitions of the SSM parameters defined for dynamical systems. This is followed by an exploration of different types of priors and prior structures, employing knowledge of the physics to inform the prior.

## 2. STOCHASTIC STATE SPACE MODEL

In general, a stochastic output-only system can be represented by the following set of LTI state space equations [2],

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + w(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + v(t)\end{aligned}\tag{1}$$

where, as functions of time  $t$ ,  $\mathbf{y}(t) \in \mathbb{R}^l$  is the output vector,  $\mathbf{x}(t) \in \mathbb{R}^d$  is the internal state vector,  $w(t) \in \mathbb{R}^d$  and  $v(t) \in \mathbb{R}^l$  are residuals, often noted as referring to the process noise and measurement noise, respectively.  $\mathbf{A} \in \mathbb{R}^{d \times d}$  is the continuous state matrix, and  $\mathbf{C}$  is the output matrix. The process and measurement noise are both assumed to be zero-mean stationary i.i.d. processes  $w(t)$  and  $v(t)$ , respectively. In practice, output measurements are obtained at discrete points in time and so it is more common to work with the discrete state space representation:

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{A}_d \mathbf{x}_k + w_k \\ \mathbf{y}_k &= \mathbf{C} \mathbf{x}_k + v_k\end{aligned}, \quad \mathbb{E} \left[ \begin{bmatrix} w_q \\ v_q \end{bmatrix} \begin{bmatrix} w_s^\top & v_s^\top \end{bmatrix} \right] = \begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^\top & \mathbf{R} \end{bmatrix} \delta_{qs}\tag{2}$$

where  $\mathbf{y}_k$  is the output vector at discrete time step  $k$ ,  $\mathbf{x}_k$  is the internal state vector,  $\mathbf{A}_d$  is the discrete state matrix such that  $\mathbf{A}_d = \expm(\mathbf{A}\Delta t)$  given  $\Delta t$  is the sampling time.  $w_k$  and  $v_k$  refer to samples of the process noise and measurement noise, respectively.  $\mathbb{E}[\cdot]$  denotes the expectation and  $\delta_{qs}$  is the Kronecker delta for any two samples in time  $q$  and  $s$ , such that the right-hand side of Eq (2) is the covariance.

The set of equations in Eq.(1) are valid for a broad range of applications including structural dynamics. Using this model, system identification algorithms attempt to learn the parameters of this general model from observed data.

### 2.1. Theoretical State Space Parameters of a Structural System

Recall that the response of a multi degree-of-freedom linear dynamic system can be written in the form,

$$\mathcal{M}\ddot{\mathbf{y}} + \mathcal{C}\dot{\mathbf{y}} + \mathcal{K}\mathbf{y} = \mathbf{f}\tag{3}$$

where  $\mathbf{y}$  is the response and  $\dot{\mathbf{y}}, \ddot{\mathbf{y}}$  its derivatives w.r.t. time,  $\mathcal{M}, \mathcal{C}, \mathcal{K}$  are the mass, damping and stiffness matrices respectively and  $\mathbf{f}$  is the forcing. Given this model, the parameters of the continuous state space will be,

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbb{I} \\ -\mathcal{M}^{-1}\mathcal{K} & -\mathcal{M}^{-1}\mathcal{C} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \mathbb{I} & \mathbf{0} \end{bmatrix}\tag{4}$$

where  $\mathbb{I}$  is the identity matrix. These parameters exactly describe the second-order system of equations in Eq.(3), assuming displacement is observed. By performing an eigen-decomposition on the state matrix, estimates for the natural frequencies, damping ratios and the mode shapes, in combination with the output matrix, can be obtained directly [12]. As perhaps expected, the modal properties extracted from the theoretical state matrix dutifully obey the physics.

It is also useful to consider form of the state space in modal coordinates. Assuming the system parameters are diagonalisable under proportional Rayleigh damping assumptions, the equations of motion can be decoupled and the system represented as a truncated sum of smaller single DOF systems. The SSM parameters then become

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbb{I} \\ -\mathbf{\Omega}^2 & -\mathbf{\Lambda} \end{bmatrix}, \quad \mathbf{C} = [\mathbf{\Phi}^T \quad \mathbf{0}] \quad (5)$$

where  $\mathbf{\Omega}^2 = \text{diag}\{\omega_1, \dots, \omega_n\}$ ,  $\mathbf{\Lambda} = \text{diag}\{2\zeta_1\omega_1, \dots, 2\zeta_n\omega_n\}$  are diagonal matrices for  $n$  distinct modes, given that  $\omega_n$  is a natural frequency and  $\zeta_n$  the corresponding damping ratio.  $\mathbf{\Phi}$  is the mode shape matrix.

Although the diagonalised state matrix representation in Eq. (5) enforces the most structure, there is, of course, no particular reason why the proportional damping condition should be satisfied. In general, a damped linear system cannot be decoupled by classic modal analysis. However, it is frequently attempted to be represented it as such.

## 2.2. Identified State Matrix

In most system identification methods, the parameters of the SSM are identified directly from data. However, most approaches are not constructed with any particular type of dynamic system in mind. The parameters are not principally constrained by construction, nor is prior information employed. Rather, these methods extract the dominant dynamics in first-order state-space form, without directly enforcing the second-order structure that exists for a structural systems, i.e.  $\mathbf{x}_{k+1} = \hat{\mathbf{A}}\mathbf{x}_k + w_k$

Consequently, the identified state matrix (denoted  $\hat{\mathbf{A}}$ ) does not display the structure shown in Eq. 4.

$$\hat{\mathbf{A}} \neq \begin{bmatrix} \mathbf{0} & \mathbb{I} \\ -\mathcal{M}^{-1}\mathcal{K} & -\mathcal{M}^{-1}\mathcal{C} \end{bmatrix} \quad (6)$$

However, the identified state matrix still encodes information on the dynamics by way of its eigenvalues (modal frequency and damping), which should be very close to that of the true system. SSM algorithms essentially estimate a transformed version of the state matrix  $\hat{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ , where  $\mathbf{T}$  is an unknown invertible change of basis such that the matrix  $\hat{\mathbf{A}}$  is *similar* (in a linear algebra sense) to the true state matrix  $\mathbf{A}$ ; its eigen structure is invariant under this change of basis. Nevertheless, as first alluded to in the introduction, this lack of enforced structure on  $\hat{\mathbf{A}}$  can result in solutions that violate key assumptions or the physics of the system in question; system identification tools are regularly performed at a higher than actual number of modes, overfitting the data and returning spurious solutions.

## 3. PRIOR OVER THE STATE MATRIX

In this section, a set of potential prior distributions over the state matrix are presented and visualised. All priors are shown using distributional estimates over the individual elements of a 2DOF state matrix. The distributions in each case are represented by histograms of 10 000 Monte Carlo samples. Distributions over the corresponding natural frequencies and damping ratios arising from each state matrix are also provided in each case.

### 3.1. Matrix-Normal

Perhaps the simplest choice of prior, and that with the least restriction on the form of the state matrix, is the Matrix-Normal:

$$\mathbf{A} \sim \mathcal{MN}_{n,p}(\mathbf{P}, \mathbf{U}, \mathbf{V}) \quad (7)$$

where  $\mathbf{P}$  is the mean of the matrix,  $\mathbf{U}$  is the row-wise covariance and  $\mathbf{V}$  is the column-wise covariance.

This proper prior provides adequate control over the mean and the covariances of the rows and columns, making it a very flexible prior over  $\mathbf{A}$ , without any enforced structure. Figure 1a shows the prior distribution over each of the elements of a 2DOF state matrix as histograms of the samples drawn from a Matrix-Normal with  $P = \mathbf{0}$ ,  $U = \mathbb{I}$ ,  $V = \mathbb{I}$ , equivalent to  $A_{i,j} \sim \mathcal{N}(0, 1)$ . The distributions over the associated natural frequencies and damping ratios are shown in Figures 2a and 3a.

In practice, Bayesian SSM system identification methods (including SSI) regularly apply this type of prior over the *discrete* state matrix [13–15], given that the system is principally determined using discrete output measurements. Recall that the discrete state matrix is related to the continuous state matrix through the following relationship;  $\mathbf{A}_d = \expm(\mathbf{A}\Delta t)$  and the modal parameters extracted in the usual way through an eigenvalue decomposition [12].

### 3.2. Physics-informed Priors

To provide a more informative choice of prior, a selection of physics-informed priors for  $\mathbf{A}$ , based on the two theoretical forms of the state matrix given in Eq.(4) and Eq.(5) are considered. In both examples, specific elements of the matrix  $\mathbf{A}$  are fixed at either zero or one depending on the structure.

#### 3.2.1. Priors over the system parameters

Assuming that  $\mathbf{A}$  can be represented by the theoretical form in (4), priors can be introduced over the system matrices  $\mathcal{M}$ ,  $\mathcal{K}$  and  $\mathcal{C}$ . The matrices  $\mathcal{M}$ ,  $\mathcal{K}$  must have the property of being positive definite to ensure a physical and stable structure [16]. Furthermore, under proportional damping assumptions, the matrix  $\mathcal{C}$  must also be positive definite. Consequently, these matrices can be modelled by an inverse-Wishart distribution, whose samples are guaranteed to be positive definite [17]

$$\mathcal{M} \sim \mathcal{IW}(\mathbf{K}_{\mathcal{M}}, \nu) \quad , \quad \mathcal{K} \sim \mathcal{IW}(\mathbf{K}_{\mathcal{K}}, \nu) \quad , \quad \mathcal{C} \sim \mathcal{IW}(\mathbf{K}_{\mathcal{C}}, \nu) \quad (8)$$

where  $\mathbf{K}_{\bullet} \in \mathbb{R}^{D \times D}$  is a scale matrix and  $\nu$  the degrees of freedom. Figure 1b shows the prior distributions over each of the elements of a 2DOF state matrix as histograms of 10 000 samples drawn from the above inverse-Wishart priors given  $\mathbf{K}_{\mathcal{M}} = \mathbb{I}$ ,  $\mathbf{K}_{\mathcal{K}} = \mathbf{K}_{\mathcal{C}} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ ,  $\nu = D + 1$ , and with a scaling factor of 1000 applied to  $\mathcal{K}$ . The corresponding natural frequency and damping ratio estimates are also provided in Figures 2b and 3b.

#### 3.2.2. Priors over Modal Properties

Alternatively, one potential prior structure that provides a deeper embedding of existing knowledge onto the state matrix, involves introducing priors over the modal properties in the diagonalised form from Eq.(5). This particular structure enforces the Rayleigh proportional damping condition very explicitly, but gives much greater control over prior belief about the natural frequencies and damping ratios. Here two possible choices of prior are provided.

### Uniform Priors

The first of these prior choices is uniform priors placed over the natural frequencies and damping ratios. Naturally, a uniform prior over the damping ratio results in a proper prior, bounded between zero and one. On the other hand, the natural frequencies are theoretically bounded zero to infinity and thus improper under a uniform prior. However, with sensible assumptions, one can limit the range of the uniform distribution to  $[a, b]$ , resulting in a proper prior.

$$\omega_n \sim \mathcal{U}(a_{\omega_n}, b_{\omega_n}) \quad , \quad \zeta_n \sim \mathcal{U}(a_{\zeta_n}, b_{\zeta_n}) \quad (9)$$

where  $\mathcal{U}(a, b)$  is a uniform distribution bounded between  $[a, b]$ . Figure 1c shows the prior distribution over the elements of the 2DOF state matrix whilst the corresponding natural frequency and damping ratio estimates are given in Figures 2c and 3c for  $\omega_{1,2} \sim \mathcal{U}(0, 20)$  and  $\zeta_{1,2} \sim \mathcal{U}(0, 1)$ .

### Log Normal and Beta Priors

The second physics-based prior sees a more specific and informative choice of prior presiding over the modal properties. The first is a Gaussian prior over the log of the natural frequencies (i.e. a lognormal). As the Gaussian distribution has support over  $\mathbb{R}$  i.e. bounded  $[-\infty, \infty]$ , potentially violating the condition of only positive natural frequency estimates, choosing a lognormal (with support over  $\mathbb{R}^+$ ) is perhaps a more appropriate choice. It is noted that when considering estimates far from  $\mathbb{R}^-$  under low variances (i.e. the central limit theorem), a Gaussian prior is an acceptable approximation. Alternatively, a gamma prior would also be a suitable choice, given its support over  $\mathbb{R}^+$ . Here the lognormal is chosen. The distribution over the damping ratios are assumed to follow a Beta distribution, which by definition is bounded between  $[0, 1]$ . These priors are defined as,

$$\ln(\omega_n) \sim \mathcal{N}(\mu_n, \sigma_n^2) \quad , \quad \zeta_n \sim \mathcal{B}(\alpha_n, \beta_n) \quad (10)$$

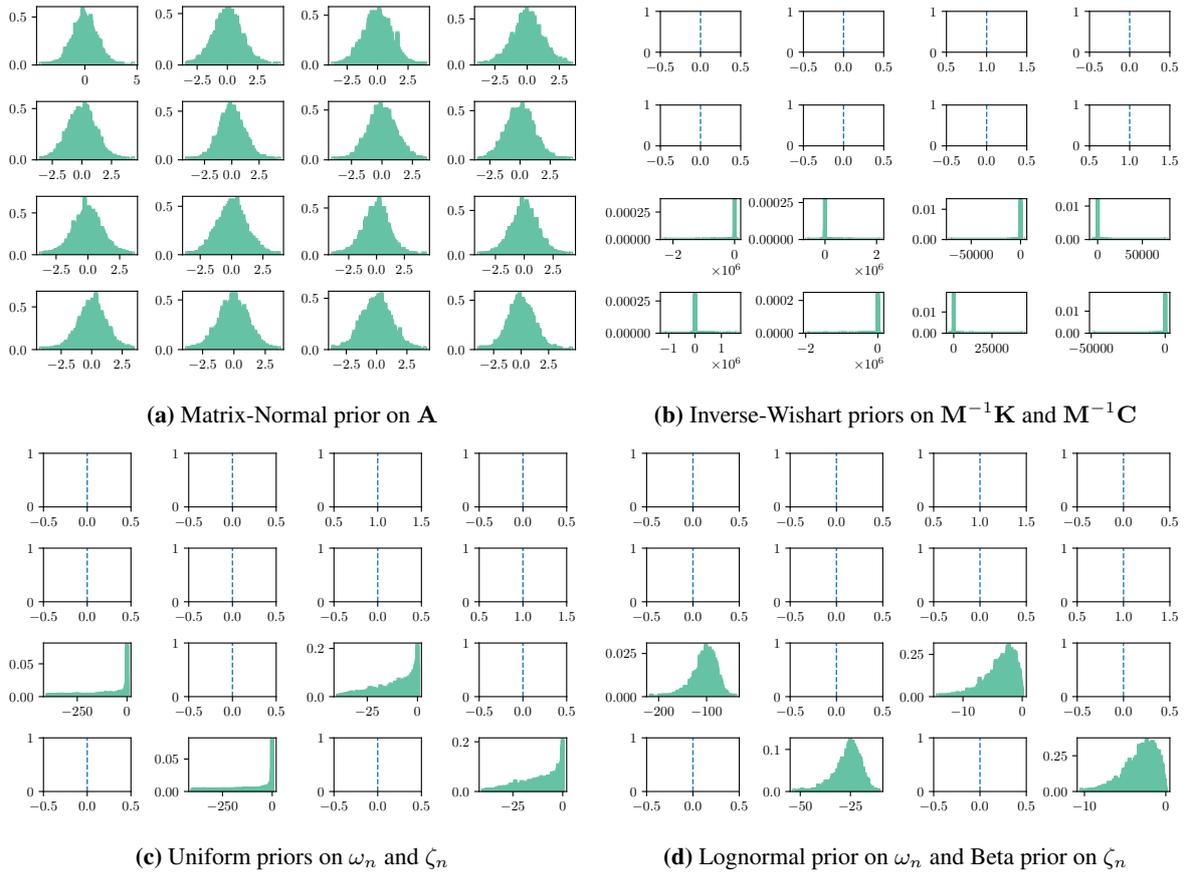
where  $\mathcal{N}(\mu, \sigma)$  denotes a Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$ , and  $\mathcal{B}(\alpha, \beta)$  denotes a Beta distribution controlled by shape parameters  $\alpha$  and  $\beta$ . The prior distributions over the elements of the 2DOF state matrix are shown in Figure 1a, given  $\ln(\omega_1) \sim \mathcal{N}(\ln(5), 0.1)$ ,  $\ln(\omega_2) \sim \mathcal{N}(\ln(10), 0.1)$ ,  $\zeta_1 \sim \mathcal{B}(2, 10)$  and  $\zeta_2 \sim \mathcal{B}(2, 5)$ . The corresponding natural frequency and damping ratio estimates are also shown in Figures 2d and 3d, respectively.

## 4. PRIOR OVER THE OUTPUT MATRIX

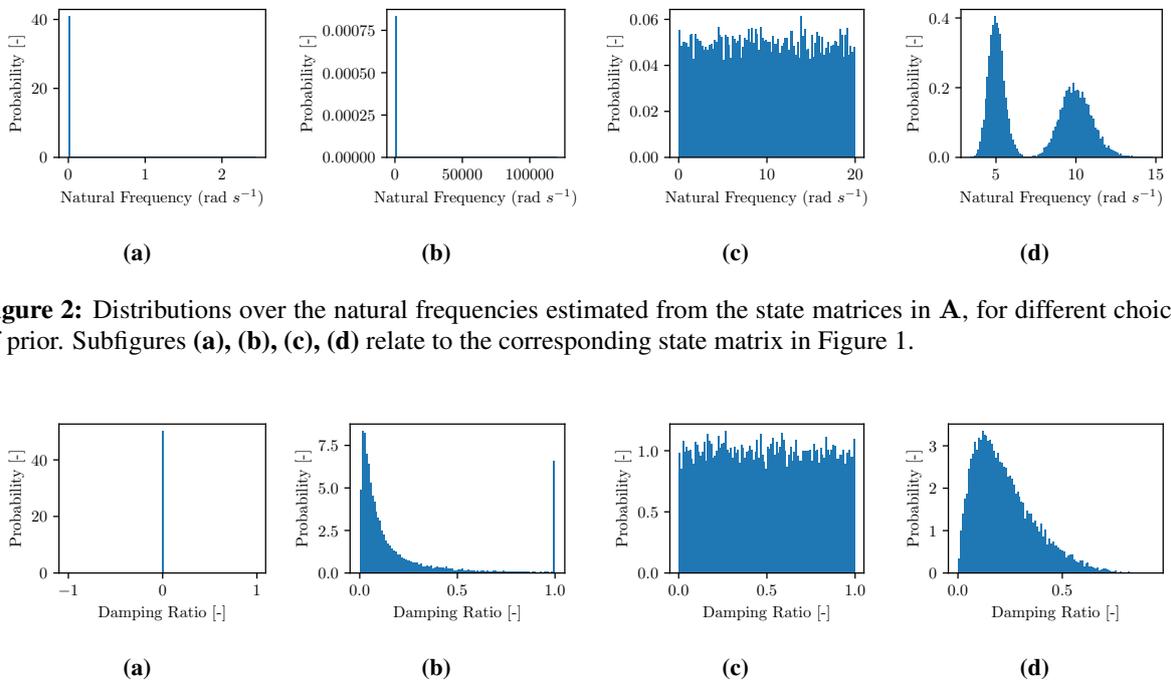
It is also useful to briefly discuss the prior over the output matrix  $\mathbf{C}$ . The output matrix is used in combination with the eigenvectors of  $\mathbf{A}$  to obtain the mode shapes [18]. If displacement is observed, then  $\mathbf{C}$  should take the form shown in Eq. (4). In contrast, if acceleration is observed, then  $\mathbf{C}$  would be the lower half of the state matrix  $\mathbf{A}$  if constrained to the form  $\mathbf{A} = \begin{bmatrix} 0 & \mathbb{I} \end{bmatrix}$  e.g. in Eq. (4) [19]. In general, the output matrix could take any form, meaning a Matrix-Normal distribution is perhaps a sensible choice.

## 5. DISCUSSION

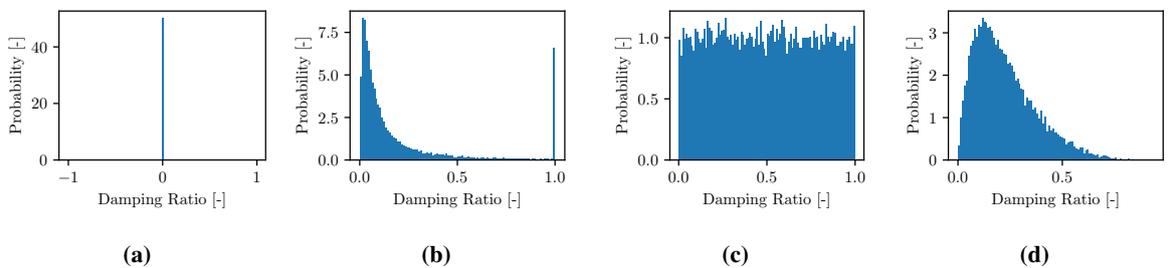
As is evident from Figures 2 and 3, the choice and specification of the prior over the state matrix, and the matrix structure (seen in Figure 1), has a significant effect on the prior over the modal properties. Each prior structure offers different guarantees as to the nature of the corresponding priors over the modal properties and crucially, determines if estimates are to be physically meaningful in the context of structural dynamic systems.



**Figure 1:** Distributional estimates over the elements of the state matrix  $\mathbf{A}$  for different choices of prior structure.



**Figure 2:** Distributions over the natural frequencies estimated from the state matrices in  $\mathbf{A}$ , for different choices of prior. Subfigures (a), (b), (c), (d) relate to the corresponding state matrix in Figure 1.



**Figure 3:** Distributions over the damping ratios estimated from the state matrices in  $\mathbf{A}$ , for different choices of prior. Subfigures (a), (b), (c), (d) relate to the corresponding state matrix in Figure 1.

The Matrix-Normal prior, although providing the greatest flexibility, offers no guarantee that estimates will be physically meaningful. Alternatively, the remaining three prior structures do guarantee physically meaningful estimates by imposing constraints on certain elements of the matrix and more complex combinations of priors. Despite this improvement, there remains an open question regarding which of these priors can provide sufficient flexibility to measured data, whilst also ensuring the desired physicality.

In the context of Bayesian inference, the more complex the choice of prior, often the harder or more computationally demanding the inference may be; especially if priors do not exhibit any conjugacy. This is one of the difficult choices when considering which prior to choose in a model. An uninformative or weakly informative priors may be chosen over a more informative prior if it can significantly simplify the inference. However, the authors note that this decision could be considered counterproductive, given the initial reason for introducing prior information.

Finally, although this work explored a possible set of priors for the state matrix  $A$ , some models require prior distributions over the observability and controllability matrices (e.g. [7]). This requirement introduces another level of abstraction above the modal properties, back through a classic linear solve [2]. How these priors could be suitably propagated onto the observability and controllability matrices is an open and interesting problem to be explored.

## 6. CONCLUSIONS

In this work, a range of possible prior distributions over the state matrix  $A$  have been explored. The need to explore such priors has arisen from an increase in Bayesian methods to uncertainty quantification in OMA and more broadly a general desire for physically meaningful estimates of the modal properties; not necessarily guaranteed using SSMs LTI system identification methods. Four different types of prior were presented, the structures of which ranged from very flexible, weakly informative priors mirroring current approaches, to those based on the physics; informing the different types of structure and distribution imposed on the elements of the state matrix. The distributions over the elements of the state matrix and the corresponding natural frequencies and damping ratios were also shown, demonstrating the variability in the prior estimates.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Engineering and Physical Sciences Research Council (EPSRC), UK through programme grant number EP/W005816/1. For the purpose of open access, the author has applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising.

## REFERENCES

- [1] L Ljung. *System Identification: Theory for the User*. Prentice-Hall Information and System Sciences Series. Prentice Hall, Upper Saddle River, NJ, USA, 2nd ed., 2012.
- [2] T Katayama. *Subspace Methods for System Identification: A Realisation Approach*. Communications and Control Engineering. Springer, Berlin, London, 2005.
- [3] T. B Schön, A Wills, and B Ninness. System identification of nonlinear state-space models. *Automatica*, 47(1):39–49, 2011.
- [4] B Peeters and G De Roeck. Reference-based stochastic subspace identification for output-only modal analysis. *Mechanical Systems and Signal Processing*, 13(6):855–878, 1999.
- [5] P Van Overschee and B De Moor. *Subspace Identification for Linear Systems*. Springer, Boston, MA, USA, 1996.

- [6] B. J O'Connell and T. J Rogers. A robust probabilistic approach to stochastic subspace identification. *Journal of Sound and Vibration*, 581:118381, 2024.
- [7] B. J O'Connell, M. D Champneys, and T. J Rogers. A new perspective on bayesian operational modal analysis. *arXiv preprint arXiv:2408.08664*, 2024.
- [8] B Li, A Der Kiureghian, and S.-K Au. A Gibbs sampling algorithm for structural modal identification under seismic excitation. *Earthquake Engineering & Structural Dynamics*, 47(14):2735–2755, 2018.
- [9] K Xu, Q.-S Li, K Zhou, and X.-L Han. A bootstrap-based stochastic subspace method for modal parameter identification and uncertainty quantification of high-rise buildings. *Journal of Building Engineering*, 87:109007, 2024.
- [10] B Li and A Der Kiureghian. Operational modal identification using variational Bayes. *Mechanical Systems and Signal Processing*, 88:377–398, 2017.
- [11] Y Huang, J. L Beck, and H Li. Bayesian system identification based on hierarchical sparse Bayesian learning and Gibbs sampling with application to structural damage assessment. *Computer Methods in Applied Mechanics and Engineering*, 318:382–411, 2017.
- [12] E Reynders. System Identification Methods for (Operational) Modal Analysis: Review and Comparison. *Archives of Computational Methods in Engineering*, 19(1):51–124, 2012.
- [13] A Wills, T. B Schön, F Lindsten, and B Ninness. Estimation of linear systems using a gibbs sampler. *IFAC Proceedings Volumes*, 45(16):203–208, 2012. 16th IFAC Symposium on System Identification.
- [14] A Svensson, A Solin, S Särkkä, and T Schön. Computationally efficient bayesian learning of gaussian process state space models. In A Gretton and C. C Robert, editors, *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*, volume 51 of *Proceedings of Machine Learning Research*, pages 213–221, Cadiz, Spain, 2016. PMLR.
- [15] A Svensson and T. B Schön. A flexible state-space model for learning nonlinear dynamical systems. *Automatica*, 80:189–199, 2017.
- [16] C Soize. Random matrix theory and non-parametric model of random uncertainties in vibration analysis. *Journal of Sound and Vibration*, 263(4):893–916, 2003.
- [17] A Gelman. *Bayesian Data Analysis*. Chapman & Hall/CRC Texts in Statistical Science. CRC Press, Boca Raton, 3rd ed., 2014.
- [18] D. J Ewins. *Modal Testing: Theory, Practice and Application*. Number 10 in Mechanical Engineering Research Studies Engineering Dynamics Series. Research Studies Press, Baldock, Hertfordshire, England, 2nd ed., 2000.
- [19] N. M. M Maia and J. M. M Silva, editors. *Theoretical and Experimental Modal Analysis*. Number 9 in Mechanical Engineering Research Studies. Research Studies Press, Taunton, Somerset, England, 1997.