White Rose Research Online

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk
Universities of Leeds, Sheffield and York

Deposited via The University of York.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/id/eprint/207117/

Version: Published Version

Article:
Yang, Jiannan (2023) Decision-Oriented Two-Parameter Fisher Information Sensitivity Using Symplectic Decomposition. Technometrics. ISSN: 1537-2723

https://doi.org/10.1080/00401706.2023.2216251

Reuse
This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:
https://creativecommons.org/licenses/

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

UNIVERSITY OF LEEDS
University of Sheffield
UNIVERSITY of York

Taylor & Francis
Taylor & Francis Group

Technometrics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/utch20

Decision-Oriented Two-Parameter Fisher Information Sensitivity Using Symplectic Decomposition

Jiannan Yang

To cite this article: Jiannan Yang (27 Jun 2023): Decision-Oriented Two-Parameter Fisher Information Sensitivity Using Symplectic Decomposition, Technometrics, DOI: 10.1080/00401706.2023.2216251

To link to this article: https://doi.org/10.1080/00401706.2023.2216251

© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.

View supplementary material

Published online: 27 Jun 2023.

Submit your article to this journal

Article views: 442

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=utch20

TECHNOMETRICS
2023, VOL. 00, NO. 0, 1–12
https://doi.org/10.1080/00401706.2023.2216251

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

Decision-Oriented Two-Parameter Fisher Information Sensitivity Using Symplectic Decomposition

Jiannan Yang[a]

Department of Engineering, University of Cambridge, Cambridge, UK

ABSTRACT

The eigenvalues and eigenvectors of the Fisher Information Matrix (FIM) can reveal the most and least sensitive directions of a system and it has wide application across science and engineering. We present a symplectic variant of the eigenvalue decomposition for the FIM and extract the sensitivity information with respect to two-parameter conjugate pairs. The symplectic approach decomposes the FIM onto an even-dimensional symplectic basis. This symplectic structure can reveal additional sensitivity information between two-parameter pairs, otherwise concealed in the orthogonal basis from the standard eigenvalue decomposition. The proposed sensitivity approach can be applied to naturally paired two-parameter distribution parameters, or a decision-oriented pairing via regrouping or re-parameterization of the FIM. It can be used in tandem with the standard eigenvalue decomposition and offer additional insights into the sensitivity analysis at negligible extra cost. Supplementary materials for this article are available online.

ARTICLE HISTORY
Received August 2022
Accepted May 2023

KEYWORDS
Conjugate parameters; Decision under uncertainty; Moment-independent sensitivity; Probabilistic sensitivity; Symplectic eigenvalue; Williamson's theorem

1. Introduction

1.1. Background

Sensitivity analysis is an integral part of mathematical modeling, and in particular a crucial element of decision-making in presence of uncertainties. The Fisher information, first introduced by Fisher and Russell (1922) and is widely used for parameter estimation and statistical inference, has found increasing application in many areas of science and engineering for probabilistic sensitivity analysis. For example, the Fisher Information Matrix (FIM) has been applied to the parametric sensitivity study of stochastic biological systems (Gunawan et al. 2005); the FIM is used to study sensitivity, robustness, and parameter identifiability in stochastic chemical kinetics models (Komorowski et al. 2011); through the link with relative entropy, the Fisher information is used to assess the most sensitive directions for climate change given a model for the present climate (Majda and Gershgorin 2010); used in conjunction with the principle of Max Entropy, the FIM is used to identify the pivotal voters that could perturb the collective voting outcomes in social systems (Lee et al. 2020); and more recently in Yang, Langley, and Andrade (2022) the Fisher information have been proposed as one of the process-tailored sensitivity metrics for engineering design. Despite the wide scope, the applications mentioned above all use the spectral analysis of the FIM, that is, the eigenvalues and eigenvectors of the FIM reveal the most and least sensitive directions of the system.

In this article, we apply a symplectic spectral analysis of the FIM, and demonstrate that the resulting symplectic eigenvalues and eigenvectors are oriented toward better decision support by extracting sensitivity information with respect to two-parameter pairs (e.g., the mean and standard deviation of a Normal distribution). As to be demonstrated, the proposed symplectic decomposition can be used in tandem with the standard eigenvalue decomposition and offer additional insights into the sensitivity analysis at negligible extra cost. An overview is given in Figure 1.

Consider a general nonlinear function $\mathbf{y} = \mathbf{h}(\mathbf{x})$: $\mathbb{R}^{\mathbf{n}} \to \mathbb{R}^{\mathbf{m}}$, the probabilistic sensitivity analysis characterizes the uncertainties of the output $\mathbf{y}$ that is induced by the random input $\mathbf{x}$ (Oakley and O'Hagan 2004; Oakley 2009). It is noted in passing that although the sensitivity analysis in this article is derivative-based, the function $\mathbf{h}(\cdot)$ needs not to be differentiable, and can be treated as black-box models when a sampling approach is used.

When the inputs can be described by parametric probability distributions (including statistically correlated inputs), that is, $x \sim p(\mathbf{x}|\mathbf{b})$, the FIM can then be estimated as the covariance matrix of the random gradient vector $\partial \ln p(\mathbf{y}|\mathbf{b})/\partial \mathbf{b}$, with the jkth entry of the FIM as (e.g., Yang, Langley, and Andrade 2022):

$$F_{jk} = \int \frac{\partial p(\mathbf{y}|\mathbf{b})}{\partial b_j} \frac{\partial p(\mathbf{y}|\mathbf{b})}{\partial b_k} \frac{1}{p} d\mathbf{y} = \mathbb{E}_Y \left[\frac{\partial \ln p}{\partial b_j} \frac{\partial \ln p(\mathbf{y}|\mathbf{b})}{\partial b_k} \right] \quad (1)$$

where $p(\mathbf{y}|\mathbf{b})$ is the joint Probability Density Function (PDF) of the outputs. The eigenvalues of the FIM represent the magnitudes of the sensitivities with respect to (wrt) simultaneous variations of the parameters $\mathbf{b}$, and the relative magnitudes and directions of the variations are given by the corresponding eigenvectors (Yang, Langley, and Andrade 2022). See Figure 1 for an overview.

CONTACT Jiannan Yang jy419@cam.ac.uk Department of Engineering, University of Cambridge, Trumpington Street, Cambridge CB2 1PZ, UK.
Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/TECH.
© 2023 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

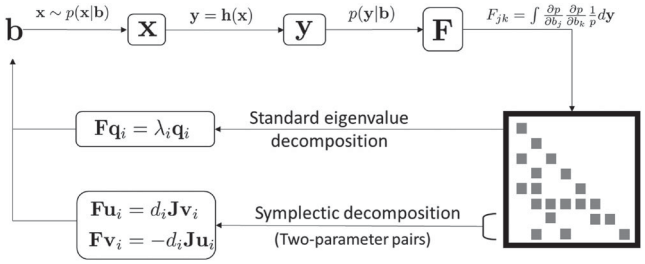


Figure 1. Overview of the sensitivity analysis based on the eigenvalues/eigenvectors of symmetric FIM (**F**). The proposed symplectic decomposition looks at the sensitivities with respect to two-parameter pairs, and it can be used in tandem with the standard approach to provide additional insights.

The FIM depends on the parameterization used. Suppose $b_j = g_j(\theta_i)$, $i = 1, 2, \ldots, s$, then the FIM with respect to the parameter $\boldsymbol{\theta}$ is (Lehmann and Casella 1998):

$$\mathbf{F}(\boldsymbol{\theta}) = \mathbb{J}^{\mathsf{T}} \mathbf{F}(\mathbf{b}) \mathbb{J} \tag{2}$$

where $\mathbb{J}$ is the Jacobian matrix with $\mathbb{J}_{ji} = \partial b_j / \partial \theta_i$. Equation (2) can be used to transform the FIM-based sensitivity analysis from a general set of distribution parameters to a re-parameterization wrt the means and standard deviations of the inputs. For example, in Section 4, an example is given for converting the sensitivities wrt the scale/shape parameters of Gamma distributions to the mean/standard deviation values of the uncertain inputs.

It should be noted that the sensitivity analysis based on FIM is fundamentally different from the commonly used variance-based analysis (Saltelli 2008). The Fisher sensitivity examines the perturbation of the entire joint PDF of the outputs, more specifically, the entropy of output uncertainty. Moreover, the sensitivity measures from the Fisher analysis are the eigenvectors, which can be regarded as the principal directions for a simultaneous variation of the input parameters. This is in contrast to variance-based ranking where it is assumed that the uncertainty of the input factors can be completely reduced to zero (Oakley and O'Hagan 2004). As pointed out in Yang (2023), using principal sensitivity directions is based on a pragmatic view that given a finite budget to change the parameters, maximizing the impact on the outputs follows the principal sensitivity directions, which tend to be a simultaneous variation of the parameters because their effects on the output are likely to be correlated. The constrained maximization view also leads to the symplectic eigenvectors in a symplectic basis as discussed in Section 2.1.

The FIM uses the information from the joint PDF of the outputs $p(\mathbf{y})$ and its gradient vector, as can be seen in (1). This is different from many of the distribution-based sensitivity analysis methods, where a distance metric is defined to measure the discrepancy between the conditional and unconditional output PDFs. For example, Borgonovo (2007) proposed a moment independent δ-indicator that looks at the entire input/output distribution. The definition of δ-indicator examines the expected total shift between the conditional and unconditional output PDFs, where the shift is conditional on one or more of the random input variables. Other distance measures have also been used for sensitivity analysis, including the mutual information (Critchfield and Willard 1986), relative entropy and the Hellinger distance (Jia and Taflanidis 2014). A review of distribution-based sensitivity methods can be found in Borgonovo and Plischke (2016).

Note that the FIM is closely linked to the relative entropy (see Section 2.2) between the jPDF of the outputs and its perturbation due to an infinitesimal variation of the input distributions. Sensitivity indices based on the modification of the input PDFs have been proposed in Lemaître et al. (2015) for reliability sensitivity analysis, where the input perturbation is derived from minimizing the probability divergence under constraints. In contrast, we consider parametric uncertain inputs in this article to form the FIM and the resulting eigenvectors provide the principal directions for the input perturbation.

The Fisher sensitivity is based on partial derivatives, but it is different from the derivative-based global sensitivity measure (Sobol' and Kucherenko 2009) which is defined as the integral of the squared derivatives of the function output. The Fisher information, on the other hand, is defined as the variance of the partial derivatives of the log probability of the uncertain function output, as seen in (1). And this differentiation is wrt the distribution parameters of the uncertain input, not wrt the uncertain variables themselves. Therefore, the Fisher sensitivity examines the impact of the perturbation of the input probability distribution, and as the input distributions are often estimated from data, it is equivalent to assessing which uncertain datasets to be focused on.

Many widely applied parametric distributions are in the two-parameter families, for example, the location-scale families including the Normal distribution and Gamma distribution. Although the Fisher sensitivity is with respect to these distribution parameters **b**, the quantities of interest for decision-making are ultimately the uncertain variables **x** themselves, for example, to rank the relative importance of **x**. We will demonstrate in this article that the symplectic decomposition of the FIM identifies the influential two-parameter pairs, or equivalently the corresponding variables, and can be used in tandem with the standard eigenvalue decomposition for better decision supports.

1.2. A Motivating Example

As a motivating example, we consider an engineering design problem under uncertainties. Consider a simple cantilever beam where the Young's modulus E and the length L are uncertain, that is, $\mathbf{x} = (E, L)$, and the uncertainties can be described by Normal distributions with $E \sim \mathcal{N}(\mu_1 = 69e9,\ \sigma_1^2 = 11.5e9^2)$ and $L \sim \mathcal{N}(\mu_2 = 0.45,\ \sigma_2^2 = 0.045^2)$. To keep it analytically tractable for this motivating example, we assume a trivial function $\mathbf{y} = \mathbf{x}$ (a random vibration problem considered in Section 4). Assuming the two random variables are independent, the FIM in this case is diagonal (Cover and Thomas 2006):

$$\mathbf{F}(\mu_1, \sigma_1, \mu_2, \sigma_2) = \operatorname{diag}\left(\sigma_1^{-2}, 2\sigma_1^{-2}, \sigma_2^{-2}, 2\sigma_2^{-2}\right). \tag{3}$$

The eigenvalues, the diagonal entries of the FIM in (3) in this case, and the corresponding eigenvectors then provide the sensitivity information of the uncertain output **y** wrt the input parameter vector $\mathbf{b} = (\mu_1, \sigma_1, \mu_2, \sigma_2)$. More specifically, they correspond to the sensitivities of the relative entropy of the random outputs as to be discussed in Section 2 (Yang, Langley, and Andrade 2022).

For practical use of the Fisher sensitivity information, two issues need to be addressed and that motivates our research in this article. First, the FIM needs to be normalized. On one hand,

the un-normalized FIM given in (3) tends to be ill-conditioned. For example, the condition number is in the order of 10^{22} given that $\sigma_1 = 11.5e9$ and $\sigma_2 = 0.045$. On the other hand, as the Young's modulus E and the length L are of different units, the FIM needs to be normalized so that the sensitivities wrt the different parameters are comparable. One option is to consider sensitivity wrt a percentage change of the parameter and this is called proportional (Yang, Langley, and Andrade 2022) or logarithmic (Pantazis, Katsoulakis, and Vlachos 2013) normalized FIM. Normalization is equivalent to a re-parameterization. In the case of proportional normalization, the change of parameter is $b_j = \bar{b}_j\theta_j$ with $\bar{b}_j$ the nominal value for normalization, and the Jacobian matrix in (2) is just a diagonal matrix with $\bar{b}_j$ on the diagonal.

However, proportional normalization might provide unrealistic sensitivity information for practical applications. For example, unless the probability distribution of the input variables is far from the real distribution, it is most likely that the change of the mean should be within one or two standard deviations. The FIM can instead be normalized by the standard deviations, which implies that the allowable ranges of the mean values are limited to a local region and it is quantified by the standard deviations. Normalizing the FIM from (3) by the standard deviations, that is, $\bar{b}_j$ equal to the corresponding σ, we have the normalized FIM as

$$\begin{aligned}\mathbf{F}_{\text{nor}} &= \text{diag}(\sigma_1^2\sigma_1^{-2}, 2\sigma_1^2\sigma_1^{-2}, \sigma_2^2\sigma_2^{-2}, 2\sigma_1^2\sigma_2^{-2}) \\ &= \text{diag}(1, 2, 1, 2) \end{aligned} \tag{4}$$

where it is evident the condition number of the normalized FIM is much smaller. However, the $\mathbf{F}_{\text{nor}}$ in (4) has repeated eigenvalues, and as a result, the corresponding eigenvectors are not unique. Although the situation with repeated eigenvalues might seem extreme, as to be seen with more examples, the eigenvalues of the normalized FIM tend to be of similar magnitudes. In other words, the sensitivity information has been compressed by normalization (in exchange for better conditioning). As we shall see, the symplectic decomposition of the FIM has a unique symplectic structure, and that tends to mitigate this issue by making the sensitivity information for different variables more distinctive (by pairing the parameters).

The second issue with the Fisher sensitivity with respect to the distribution parameters is the gap to decision-making. The purpose of the sensitivity analysis is to identify the influential variables so that informed decisions can be made. Although it is possible to make changes to the means and standard deviations independently, the quantities of interest are ultimately the variables themselves, that is, E and L in this case. As to be demonstrated, the symplectic approach would naturally put parameters in pairs, for example, (μ, σ) as a conjugate pair for random inputs with Normal distributions, and provide more direct support for decision-making.

It is noted in passing that even if the true distribution of the uncertain input is not Normal, a common practice is still to use mean and standard deviation as the summary statistic for the dataset at hand. As a result, the two-parameter pair sensitivity proposed in this article still applies. For example, in Section 4, the sensitivities wrt the distribution parameters of Gamma random variables are re-parameterized to the means and stand deviations of the input data using (2).

1.3. Summary and Paper Outline

In summary, the use of the Fisher information as a sensitivity measure has wide applications across science and engineering. Nevertheless, practical issues can hinder the translation of sensitivity information into actionable decisions. In this article, we propose a new approach using the symplectic decomposition to extract the Fisher sensitivity information. The symplectic decomposition uses Williamson's theorem (Williamson 1936; Nicacio 2021) which is a key theorem in Gaussian quantum information theory (Pereira, Banchi, and Pirandola 2021). Originating from Hamiltonian mechanics, the symplectic transformations preserve Hamilton's equations in phase space (Arnol'd 1989). In analogy to the conjugate coordinates for the phase space, that is, position and momentum, we regard the input parameters as conjugate pairs and use a symplectic matrix for the decomposition of the FIM. The resulting symplectic eigenvalues of large magnitudes, and the corresponding symplectic eigenvectors of the FIM, then reveal the most sensitive two-parameter pairs.

It should be noted that the proposed symplectic decomposition is only applicable for parameter space of even dimensions, that is, $\mathbf{b} \in \mathbb{R}^{2n}$, and the requirement that the parameters can be regarded as two-parameter pairs. For the two-parameter family of probability distributions, such as the widely used location-scale families, there is a natural pairing of the parameters. For other cases, a decision-oriented pairing might be needed. For example, a re-parameterization wrt the mean and standard deviation, or two moments of the random variables, using (2) would transform the FIM into even dimensions. Once the FIM is obtained wrt parameters of even dimensions, it is envisaged that the proposed symplectic decomposition is best used in tandem with the standard eigenvalue decomposition for sensitivity analysis using the FIM. This offers additional insight into the sensitivity analysis, and as the main computational burden is often at estimating the FIM, at negligible extra cost.

In what follows, we will first review the approach of symplectic decomposition using the Williamson's theorem in Section 2. The details of finding the symplectic eigenvalues and eigenvectors are given in Section C of the supplementary material together with the corresponding Matlab script. We then give a theoretical comparison between the symplectic decomposition and the standard eigenvalue decomposition, in terms of the sensitivity of entropy and also from an optimization point of view using trace maximization. A benchmark study is conducted in Section 3, where the similarity and difference between the Fisher sensitivity and the main-effect indices used in variance-based analysis are discussed. In Section 4, a numerical example using a simple cantilever beam is used to demonstrate the effect of symplectic decomposition. Concluding remarks are given in Section 5.

2. Symplectic Sensitivity of Entropy

2.1. Symplectic Decomposition

From elementary linear algebra, we know that a real symmetric matrix $\mathbf{F}$ can be diagonalized by orthogonal matrices:

$$\mathbf{Q}^{-1}\mathbf{F}\mathbf{Q} = \Lambda \tag{5}$$

where $\mathbf{Q}$ is the orthogonal eigenvector matrix, that is, $\mathbf{Q}^{\mathsf{T}} = \mathbf{Q}^{-1}$, and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots)$ contains the real eigenvalues. And the solution to (5) can be solved using the standard eigenvalue equation:

$$\mathbf{FQ} = \mathbf{Q}\Lambda, \quad \text{with} \quad \det(\mathbf{F} - \lambda\mathbf{I}) = \mathbf{0}. \tag{6}$$

The Williamson's theorem provides us with a symplectic variant of the results above. Let $\mathbf{F} \in \mathbb{R}^{\mathbf{2n}\times\mathbf{2n}}$ be a symmetric and positive definite matrix, the Williamson's theorem says that $\mathbf{F}$ can be diagonalized using symplectic matrices (Gosson 2006; Nicacio 2021):

$$\mathbf{S}^{\mathsf{T}}\mathbf{F}\mathbf{S} = \hat{D} = \begin{bmatrix} \mathbf{D} & \\ & \mathbf{D} \end{bmatrix} \tag{7}$$

where $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_n)$ is a diagonal matrix with positive entries (d_j maybe zero if $\mathbf{F}$ is semidefinite). The d_j, $j = 1, 2, \dots, n$ are said to be the symplectic eigenvalues of matrix $\mathbf{F}$ (Bhatia and Jain 2015) and are in general not equal to the eigenvalues given in (5). The matrix $\mathbf{S} = [\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_n]$ is a real symplectic matrix, and it is called the symplectic eigenvector matrix of $\mathbf{F}$. Each symplectic eigenvalue d_j corresponds to a pair of eigenvectors $\mathbf{u}_j$, $\mathbf{v}_j \in \mathbb{R}^{2n}$:

$$\mathbf{F}\mathbf{u}_j = d_j\mathbf{J}\mathbf{v}_j; \quad \mathbf{F}\mathbf{v}_j = -d_j\mathbf{J}\mathbf{u}_j. \tag{8}$$

These eigenvector pairs can be normalized so that they form an orthonormal basis for the symplectic vector space:

$$\mathbf{u}_i^{\mathsf{T}}\mathbf{J}\mathbf{v}_j = \delta_{ij}, \quad \text{for } i, j = 1, 2, \dots, n. \tag{9}$$

Details of the symplectic decomposition and some important properties of the symplectic eigenvector matrix are given in Section A of the supplementary material.

2.2. Sensitivity of Entropy

The procedure given in the previous section tells us that there exist symplectic matrices that can decompose the FIM. As shown in Section B of the supplementary material, the standard and the symplectic eigenvectors provide the directions to maximize the matrix trace in an orthogonal and a symplectic basis, respectively, and the corresponding eigenvalues indicate the sensitivities. A special case with the FIM links its eigenvalues to the sensitivities of the Kullback-Leibler (K-L) divergence, aka relative entropy, and that is discussed in this section.

As mentioned in the introduction, consider a general function $\mathbf{y} = \mathbf{h}(\mathbf{x})$, the probabilistic sensitivity analysis characterize the uncertainties of the output $\mathbf{y}$ that is induced by the random input $\mathbf{x}$. When the joint probability distribution of the output is known, the entropy of the uncertainty can be estimated as (Cover and Thomas 2006):

$$H = -\int p(\mathbf{y}|\mathbf{b}) \ln p(\mathbf{y}|\mathbf{b}) d\mathbf{y}. \tag{10}$$

The perturbation of the entropy, defined as a relative entropy quantified using the K-L divergence, can be approximated by a quadratic form using the FIM (Yang, Langley, and Andrade 2022):

$$\begin{aligned} \Delta H &\equiv \text{KL}\left[p(\mathbf{y}|\mathbf{b})||p(\mathbf{y}|\mathbf{b}+\Delta\mathbf{b})\right] \\ &= \int p(\mathbf{y}|\mathbf{b}) \ln \frac{p(\mathbf{y}|\mathbf{b})}{p(\mathbf{y}|\mathbf{b}+\Delta\mathbf{b})} d\mathbf{y} \approx \frac{1}{2}\Delta\mathbf{b}^{\mathsf{T}}\mathbf{F}\Delta\mathbf{b} \end{aligned} \tag{11}$$

where the perturbed probability is approximated using its second order Taylor expansion (e.g., see the appendix of Yang 2022). It is noted in passing that, even without the quadratic approximation to entropy, the Fisher information can be used to quantify the distribution perturbation in its own right (Gauchy et al. 2022).

Consider the standard eigenvalue decomposition of the FIM and substitute (5) into the expression for the relative entropy in (11):

$$\begin{aligned} 2\Delta H &= \Delta\mathbf{b}^{\mathsf{T}}\mathbf{F}\Delta\mathbf{b} = (\mathbf{Q}^{-1}\Delta\mathbf{b})^{\mathsf{T}}\mathbf{\Lambda}(\mathbf{Q}^{-1}\Delta\mathbf{b}) \\ &= \sum_j^{2n} \lambda_j\xi_j^2 = \sum_j^{n} \lambda_j\xi_j^2 + \lambda_{n+j}\xi_{n+j}^2 \end{aligned} \tag{12}$$

where $\xi_j = (\mathbf{Q}^{-1}\Delta\mathbf{b})_j$ and it is clear that the eigenvalues λ_j indicate the magnitude of the entropy sensitivity.

It can be seen from (12) that the relative entropy in this quadratic form can be regarded as an ellipsoid geometrically, that is, $\sum \lambda_j\xi_j^2 = 1$. This is a consequence of the semi-positive definiteness of the FIM and the ellipsoid is proper when the FIM is positive-definite. The eigenvectors of the FIM define the principal axes and the inverse of the square roots of the corresponding eigenvalues, that is, $1/\sqrt{\lambda_j}$, are the principal radii of the ellipse. Since the principal axes are orthogonal to each other, there is no direct relationship between any pair of coordinates, say (ξ_j, ξ_{n+j}), even they are dominated by the two-parameter pairs for the same variable of interest as discussed in the introduction.

Similarly, the relative entropy in the symplectic basis can be expressed as

$$\begin{aligned} 2\Delta H &= \Delta\mathbf{b}^{\mathsf{T}}\mathbf{F}\Delta\mathbf{b} = (\mathbf{S}^{-1}\Delta\mathbf{b})^{\mathsf{T}}\hat{\mathbf{D}}(\mathbf{S}^{-1}\Delta\mathbf{b}) \\ &= \begin{bmatrix}\boldsymbol{\alpha}^{\mathsf{T}} & \boldsymbol{\beta}^{\mathsf{T}}\end{bmatrix}\begin{bmatrix}\mathbf{D} & \\ & \mathbf{D}\end{bmatrix}\begin{bmatrix}\boldsymbol{\alpha} \\ \boldsymbol{\beta}\end{bmatrix} = \sum_j^n d_j(\alpha_j^2 + \beta_j^2) \end{aligned} \tag{13}$$

where $\alpha_j = (\mathbf{S}^{-1}\Delta\mathbf{b}))_j$ and $\beta_j = (\mathbf{S}^{-1}\Delta\mathbf{b}))_{j+n}$. In contrast to (12), it can be seen that the coordinate pair (α_j, β_j) is now forced to form a circle with radius $1/\sqrt{d_j}$. The consequence is that if (α_j, β_j) corresponds to the two-parameter pairs of interest, they are symplectically equivalent, in analogy to the conjugate pair, position and momentum, in Hamiltonian mechanics.

3. Benchmark Study

The Fisher information has been introduced in Yang, Langley, and Andrade (2022) for sensitivity analysis with respect to distribution parameters. A benchmark study for Fisher sensitivity, using a linear function with decreasing coefficients and a product function with constant coefficients, has been conducted in Yang (2023). In this section, we apply the Fisher sensitivity analysis, using both standard eigenvalue decomposition and the proposed symplectic decomposition, to a high-dimensional function:

$$f(\mathbf{x}) = \mathbf{a}_1^{\mathsf{T}}\mathbf{x} + \mathbf{a}_2^{\mathsf{T}}\sin(\mathbf{x}) + \mathbf{a}_3^{\mathsf{T}}\cos(\mathbf{x}) + \mathbf{x}^{\mathsf{T}}\mathbf{M}\mathbf{x}. \tag{14}$$

This function has a 15-dimensional input vector $\mathbf{x}$ and has been used in Oakley and O'Hagan (2004) for variance-based

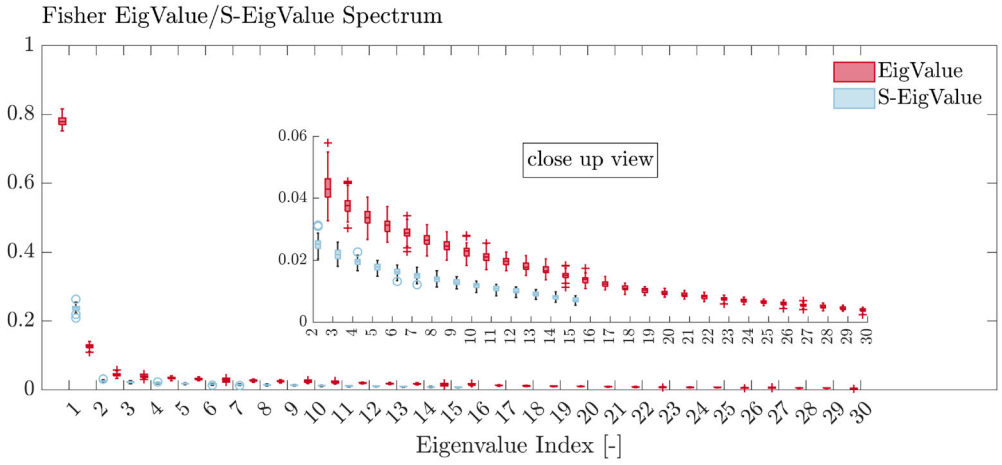


Figure 2. The standard eigenvalue and the symplectic eigenvalue (S-Eig) spectra of the FIM , for the sensitivity of the benchmark function in (14). Note that the dimension of symplectic spectrum is 15, which is half of the size of the standard eigenvalue spectrum. Results are from 50 repetitions of MC simulations, with 20,000 samples for each run. "+" and "o" indicate the outliers.

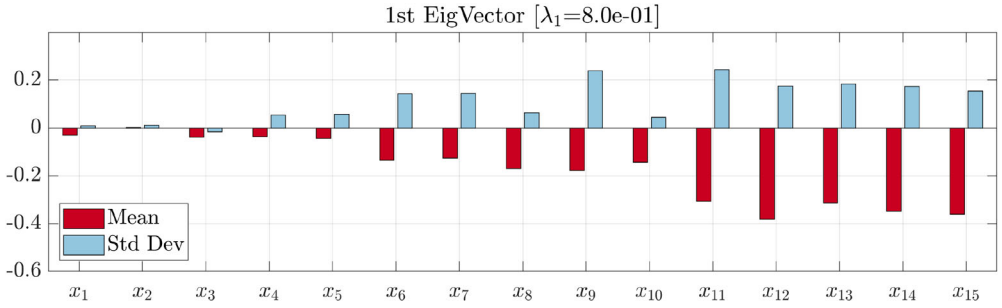


Figure 3. The first eigenvector of the FIM for the benchmark function in (14) with respect to the mean and standard deviation (Std Dev) of the 15 Gaussian inputs.

sensitivity analysis. This function's coefficients, $\mathbf{a}_1, \mathbf{a}_2$, and $\mathbf{a}_3$, are chosen so that the first five input variables have almost no effect on the output variance, x_6 to x_{10} have a much larger effect, and the remaining five contribute significantly to the output variance. All input variables are assumed to be independent and from a standard Gaussian distribution, that is, $x \sim \mathbb{N}(0, 1)$.

To estimate the FIM in (1), an efficient numerical method based on Monte Carlo sampling and the Likelihood Ratio/Score Function (LR/SF) method is used here (Rubinstein and Kroese 2016). The LR/SF method obtains a gradient estimation of a performance measure wrt continuous parameters in a single simulation run. More details of the method for FIM estimation can be found in Yang, Langley, and Andrade (2022) and Yang (2023). As the numerical computation is based on the LR/SF method, the sensitivity estimation is independent of the dimension of the input parameters (Rubinstein and Kroese 2016). This is in contrast to the variance-based methods where the computational cost is proportional to the input dimension (Saltelli 2008).

The results from the standard eigenvalue analysis of the FIM are shown in Figures 2 and 3. The eigenvalue spectrum has been computed using 20,000 Monte Carlo samples, and repeated 50 times. The bottom and top of each box in Figure 2 are the 25th and 75th percentiles of the 50 samples, respectively. Outliers, as marked as "+ and "o," are values that are more than 1.5 times the interquartile range away from the bottom or top of the box.

Only the first standard eigenvector is shown in Figure 3 as the eigenvalues corresponding to the rest of the eigenvectors are of much smaller amplitudes as seen in Figure 2. The sensitivity to the mean parameters of the input variables in Figure 3 indicates that there are three groups of importance, x_{11} to x_{15} being the most important and x_1 to x_5 being the least important. This is in good agreement with Oakley and O'Hagan (2004) from a variance-based sensitivity analysis. The sensitivity to the standard deviations, on the other hand, does not show a clear clustered trend, although it is clear that the first few variables have almost no effect.

Different from the variance-based analysis where only the amplitudes of the importance are measured, the Fisher sensitivity vectors also provide the relative phases of the sensitivity to the distribution parameters. For example, in Figure 3, it is clear that the effects of the input mean parameters on the output PDF uncertainty are in opposite directions to the effects due to the perturbation of the standard deviations. Note that the absolute sign of the eigenvector is arbitrary.

The symplectic sensitivity results are also shown in Figures 2 and 4. Different from the standard eigenvalue results, the symplectic eigenvalue spectrum has a dimension half of the standard one but the symplectic eigenvectors always come in pairs.

The symplectic sensitivity results in Figure 4 present a similar picture as the standard results, especially that the sensitivity to the standard deviations for **u1** vector indicates clear

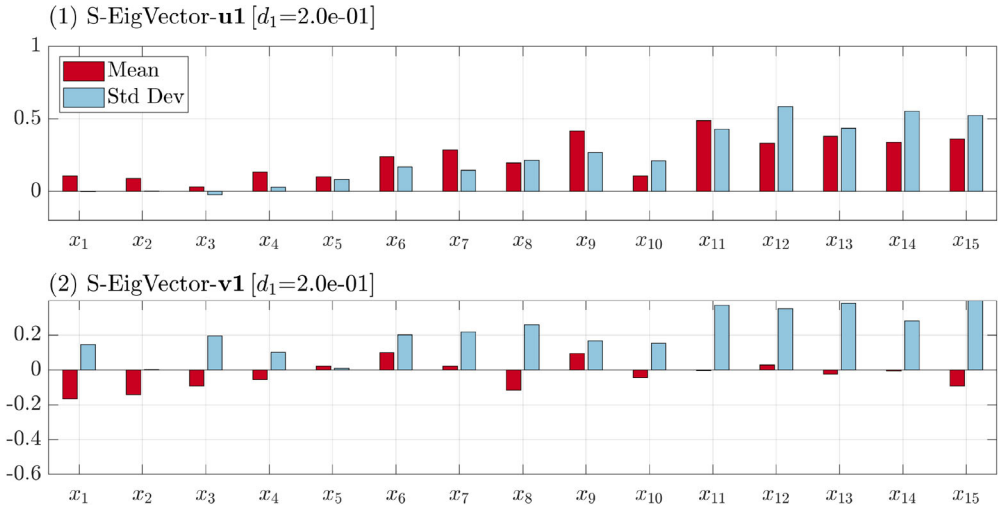


Figure 4. Same as Figure 3, but for the first pair of symplectic eigenvectors (S-EigVector) of the FIM. Note that symplectic eigenvectors come in pairs.

group importance as discussed above. However, in this case, the symplectic results do not provide any new insights. This is because the FIM is dominated by its first eigenvector in this case. For this benchmark function, there is no normalization required for the sensitivity analysis as all input variables are equivalent. As a result, there is no compression of the sensitivity information as discussed in the motivating example and the engineering example to be studied in Section 4.

The sensitivity vectors from the FIM provide the principal directions for a simultaneous variation of the input parameters. To look at the effect of individual parameters, (12) and (13) can be used. For example, assuming only parameter b_k is varied, from (12) and (13):

$$2\Delta H(\Delta b_k) = \sum_j^{2n} \lambda_j (q_{kj}\Delta b_k)^2 = \left[\sum_j^{2n} \lambda_j q_{kj}^2\right] \Delta b_k^2 \tag{15a}$$

$$2\Delta H(\Delta b_k) = \sum_j^{n} d_j (u_{kj}\Delta b_k)^2 + d_j (v_{kj}\Delta b_k)^2$$

$$= \left[\sum_j^{n} d_j \left(u_{kj}^2 + v_{kj}^2\right)\right] \Delta b_k^2 \tag{15b}$$

where q_{kj} is the kth element of the standard eigenvector $\mathbf{q}_j$, while $(\mathbf{u}_j, \mathbf{v}_j)$ are the symplectic eigenvectors. The term inside the square bracket can be regarded as the contribution to the entropy change due to the perturbation of the parameter b_k alone.

The contributions from the mean and the standard deviation parameters of the same variables can be further aggregated for the corresponding variables, assuming the perturbations are independent. For example, if (b_j, b_k) are the mean and standard deviation of the variable x_m, then the sensitivity to the variable x_m can be obtained by adding the contributions of the parameters b_j and b_k.

The resulting relative importance of the variables from the Fisher analysis, using the dominant first eigenvector (Fisher Eig) and the first pair of symplectic eigenvector (Fisher S-Eig), can then be compared to the variance based main effects (Oakley and O'Hagan 2004) and the comparison is shown in Figure 5. Although there are small deviations, the relative importance of the three groups of variables and the order of difference are clearly identified from the Fisher sensitivity analysis. Furthermore, the ratio of the first eigenvalue to the sum of all eigenvalues, as seen in Figure 2, is about 0.75 in this case and that can be regarded as the contribution of the first eigenvector to the entropy change. Although not directly comparable, 0.75 is similar to the 72% main effects contribution to the output variance as reported in Oakley and O'Hagan (2004). This offers a plausible suggestion that, in this case, the dominant first eigenvector of the FIM corresponds to the main effect from variance based sensitivity analysis.

It should be noted although the contributions from the perturbation of individual parameters are useful for benchmarking against variance-based main effects, the purpose of the Fisher sensitivity analysis is to look at the simultaneous variations of the input parameters. In contrast to variance-based analysis, the eigenvectors and symplectic eigenvectors of FIM provide principal sensitivity directions based on the impact on the joint PDF of the outputs. Not only do the eigenvectors indicate the relative amplitude, they also provide the relative phase information of the input parameter variations as discussed earlier.

In addition to the results based on 20,000 samples, Figure 5 also present the variable ranking using 1000, 5000, and 10,000 samples. Although the variability of the estimation is relatively large for smaller number of samples, the three groups of importance for the input variables are clearly identified from the FIM based indices, even with only 1000 samples for this 15-dimension problem. Note that the results in Figure 5 are normalized for comparision, and the raw data are provided in Section D of the supplementary material.

In the next section, we will consider an engineering example where the input variables are typically of different units. As their values tend to be of different orders of magnitude, normalization is required for the Fisher sensitivity analysis. In addition, different from the scalar output from this benchmark function, engineering problems tend to have multiple outputs as in the example given below.

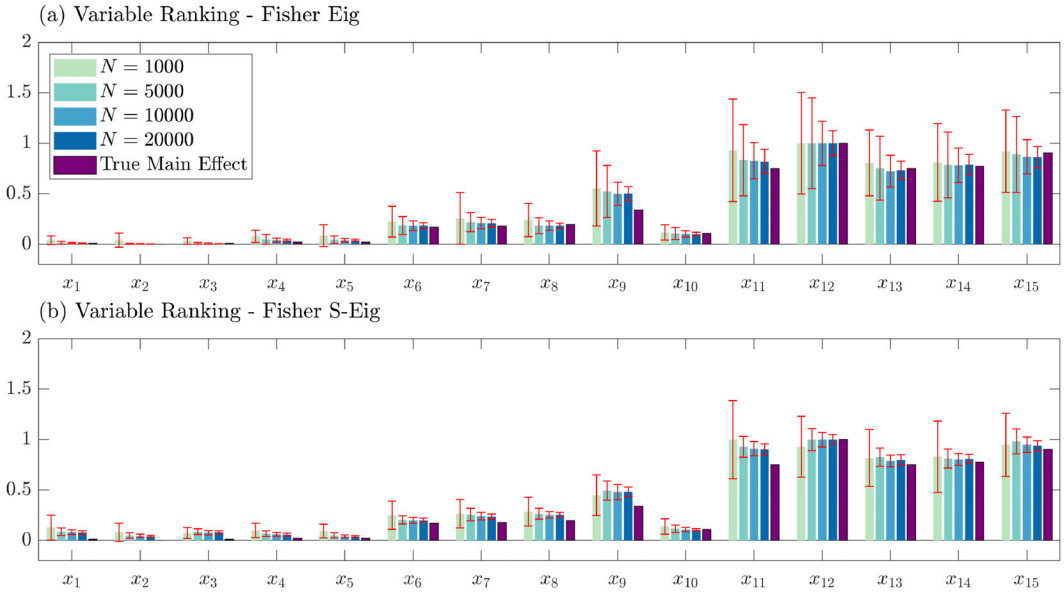


Figure 5. Variable importance ranking using three different indices: (a) the first standard eigenvector (Eig); (b) the first set of symplectic eigenvectors (S-Eig); and the true main-effect indices given in Oakley and O'Hagan (2004). The results are normalized by the largest value, where *N* indicates the number of samples used. The error bars indicate the standard deviations of the estimated Fisher indices (± one standard deviation from the mean), from repetition of 50 simulation runs.

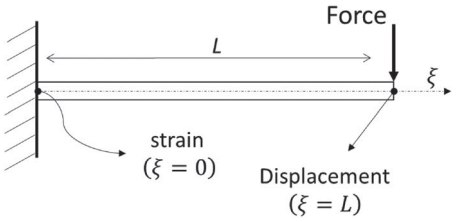


Figure 6. A cantilever beam subject to white noise excitation of unit amplitude; the responses consist of peak r.m.s displacement and strain responses.

Table 1. Mean (μ) and Coefficient of Variation (CoV) for the random variables.

	Young's Modulus	Density	Length	Width	Thickness
	$E[Pa]$	$\rho[kg/m^3]$	$L[m]$	$w[m]$	$t[m]$
μ (mean)	69e9	2700	0.45	2e-2	2e-3
	Case-1: Normal distribution				
σ/μ (CoV)	1/200	1/80	1/100	1/60	1/80
	Case-2: Gamma distribution				
σ/μ (CoV)	1/5	1/4	1/10	1/6	1/8

NOTE: Case-1 assumes Normal distribution with small CoVs, while Gamma distributions with large CoVs are considered in Case-2.

4. Application to An Engineering Example

In this section, we consider an engineering design example where the Fisher information is used for parametric sensitivity analysis of a cantilever beam. The beam is subject to a white noise excitation at the tip, see Figure 6, where the excitation is band-limited and only the first three modes are excited. In this case, the quantities of interest are the peak r.m.s responses, that is, the maximum response along the beam, for both displacement and strain (output $\mathbf{y}$ in (1) is two-dimensional). The frequency response functions for both displacement and strain responses, at different positions along the beam, are obtained via modal summation and the modal damping is assumed to be 0.1 for all modes, see Section E of the supplementary material for more details.

It is assumed that the five input variables are random, $\mathbf{x} = (E, \rho, L, w, t)$, and the sensitivities of interest are wrt their means and standard deviations, that is, $\mathbf{b} = (\mu_m, \sigma_m)$, $m = 1, 2, \ldots, 5$, as listed in Table 1. While the proposed method applies to correlated inputs, for simplicity, the inputs are independently sampled for the numerical examples considered.

Two different cases are considered. Case-1 considers Normally distributed inputs with with small variances, while Case-2 has the input variables described by Gamma distributions with larger variances. The mean values are the same for both cases.

For Normally distributed inputs, the means and standard deviations are the distribution parameters. For the inputs with Gamma distributions in Case-2, a reparameterization using (2) is required. As the mean values of our engineering example are positive, the scale (θ) and shape (k) parameters can be expressed as functions of the means and standard deviations: $k = \mu^2/\sigma^2$ and $\theta = \sigma^2/\mu$. The partial derivatives of (θ, k) wrt (μ,σ) are:

$$\frac{\partial\theta}{\partial\mu} = -\frac{\sigma^2}{\mu^2}, \quad \frac{\partial\theta}{\partial\sigma} = \frac{2\sigma}{\mu}, \quad \frac{\partial k}{\partial\mu} = \frac{2\mu}{\sigma}, \quad \frac{\partial k}{\partial\sigma} = -\frac{2\mu^2}{\sigma^3} \tag{16}$$

and these derivatives can be used to form the Jacobian matrix in (2) to re-parameterize the FIM wrt the means and standard deviations of the inputs with Gamma distributions.

It is noted in passing that it is possible to look at directly the symplectic sensitivities to the scale/shape parameters of the Gamma inputs, as they are already two-dimensional. However, we re-parameterize the FIM to demonstrate the possibility of applying the proposed symplectic approach to the mean and standard deviation parameters of non-Gaussian distributions.

For the numerical results below, the FIM is normalized by the standard deviations:

$$\mathbf{F}_{\text{nor}} = \sigma_m \sigma_n \mathbf{F}_{jk} \tag{17}$$

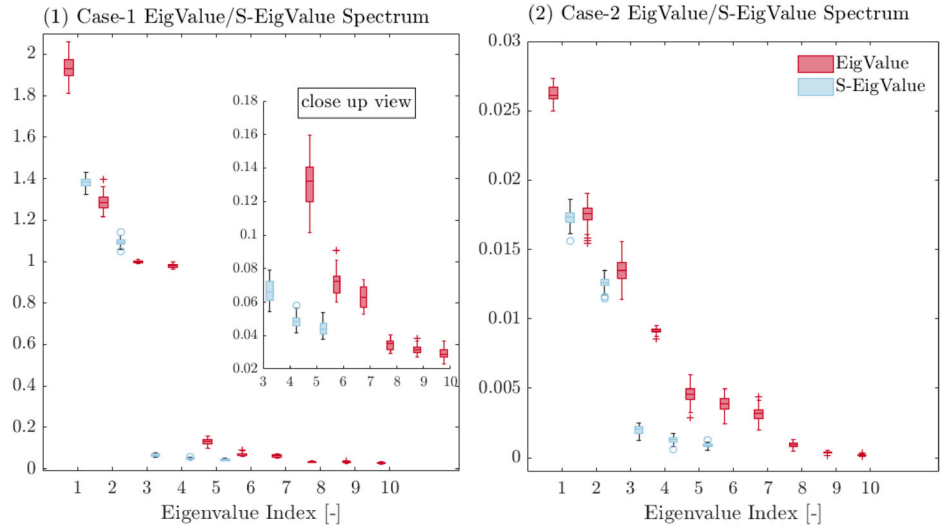


Figure 7. Eigenvalues (Eig) and Symplectic eigenvalues (S-Eig) of the FIM for (1) Case-1 ; (2) Case-2

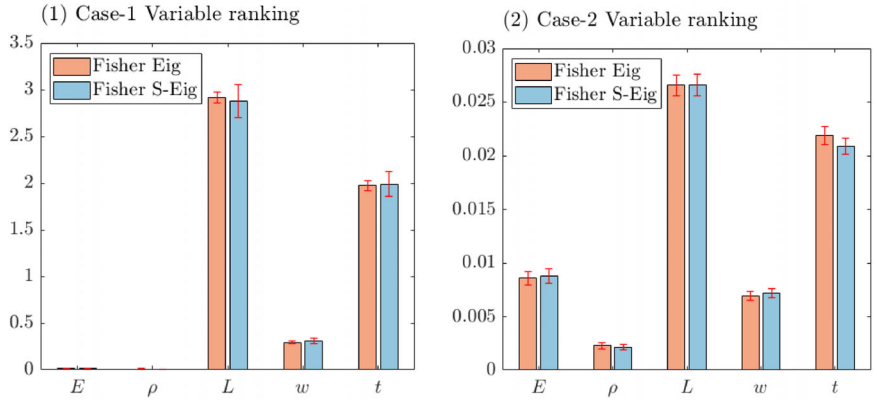


Figure 8. Overall variable ranking using the FIM for (1) Case-1 ; (2) Case-2. Error bars indicate the standard deviation from repetitions of the FIM estimation.

where $j, k = 1, 2, \ldots, 10$ and $m = j/2$, $n = k/2$ when j, k are even, and $m = (j+1)/2$, $n = (k+1)/2$ when j, k are odd numbers. As discussed in the introduction, the normalization is necessary for practical applications as the input variables are of different units and often differ by orders of magnitude. Moreover, it largely improves the condition number of the FIM. For example, in this case study, the condition number of the FIM in the order of $10^{27} \sim 10^{31}$ for both Case-1 and Case-2, and it reduces to the order of 10^2 for both cases after normalization.

Once the FIM is estimated and normalized, the standard approach is to compute the eigenvalues and eigenvectors of the FIM for sensitivity analysis as in Yang, Langley, and Andrade (2022) and Yang (2023). As discussed, the eigenvalues of the FIM represent the magnitudes of the sensitivities with respect to simultaneous parameter variations, and the most sensitive directions are given by the eigenvectors corresponding to the largest eigenvalues.

The standard eigenvalues and eigenvectors are denoted as "EigValue" and "EigVector" and are shown in Figures 7, 9(a), 10(a). The aggregated indices using (15) are also displayed in Figure 8, using the first four dominant eigenvectors (Fisher Eig), corresponding to those shown in Figures 9 and 10. The overall variable rankings between the standard and the symplectic approaches, as shown in Figure 8, are almost the same. This is as expected as they are from the same FIM. However, as to be seen, the symplectically decomposed eigenvectors can reveal additional insights into the parameter sensitivities.

The results shown in this section are based on repeated estimations of the FIM with 20,000 samples for each run. Case-1 results are obtained from 30 repeated runs, while 50 repetitions are conducted for Case-2. The eigenvectors shown in this section, as in Figures 9–11, are based on the averaged FIM. The variations of the eigenvectors are not shown as the eigenvectors can have arbitrary signs. More importantly, the symplectic eigenvectors come in pairs, which makes it difficult to compare individual vectors. Nevertheless, from Figures 7 and 8, it can be seen that the variability of the results presented in this section is reasonably low.

In Figure 7, the eigenvalues are ordered from large to small, with the first four larger than the rest, especially for Case-1. Note that the spectrum here is quite different from the benchmark case shown in Figure 2 where only one eigenvalue dominates. The corresponding first four eigenvectors are displayed in Figures 9 and 10.

As the main purpose is to compare the symplectic analysis against the standard decompositions, we will not go into details of the relative importance of the parameters. From the eigenvector results in Figures 9(a) and 10(a), an important feature is that,

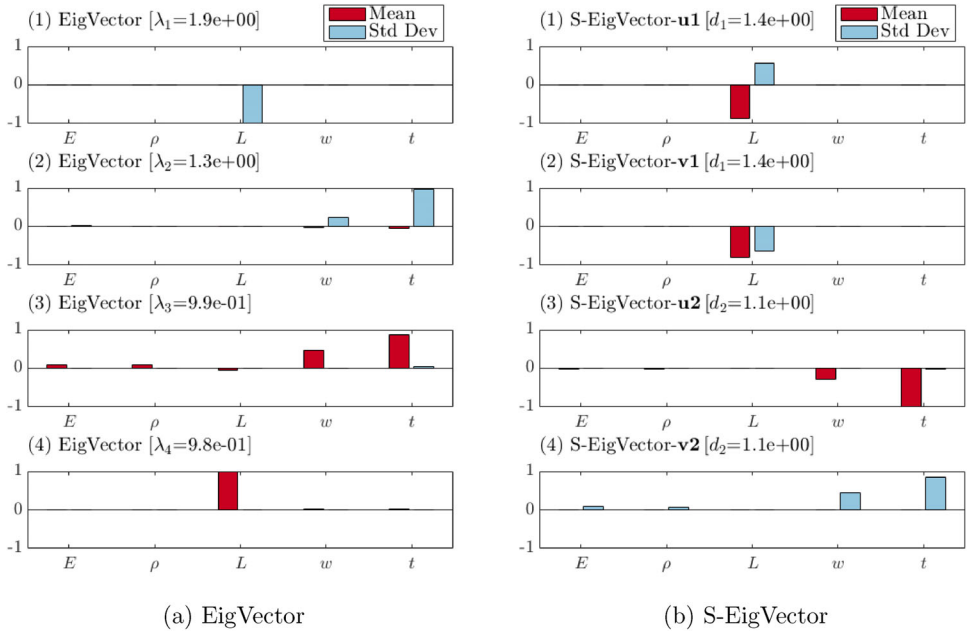


Figure 9. Eigenvectors (EigVector) and symplectic eigenvectors (S-EigVector) of the FIM for Case-1. The symplectic eigenvectors come in pairs, $\mathbf{u}_1, \mathbf{v}_1$ and $\mathbf{u}_2, \mathbf{v}_2$, and each pair corresponds to the same symplectic eigenvalue

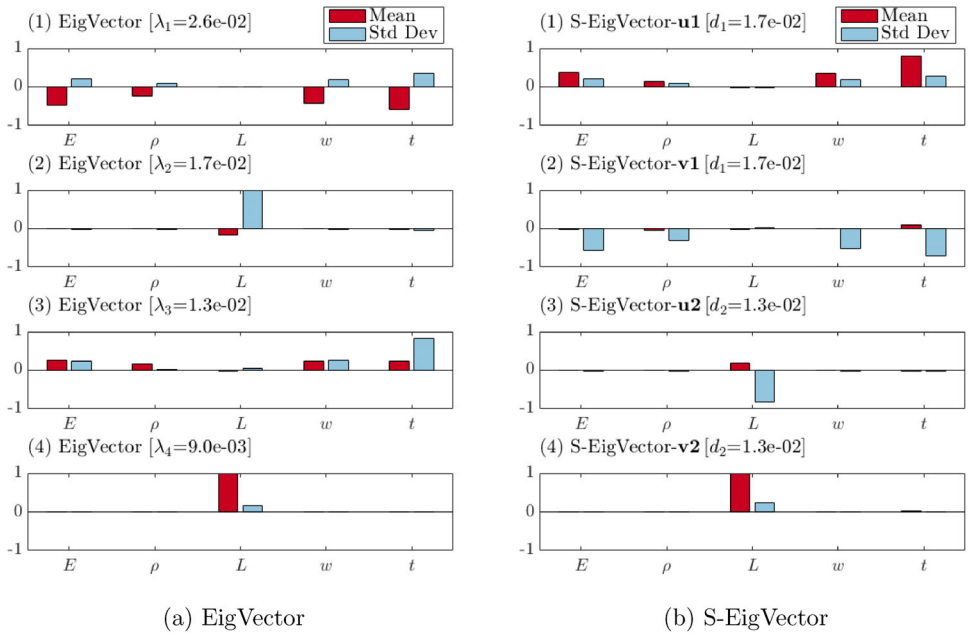
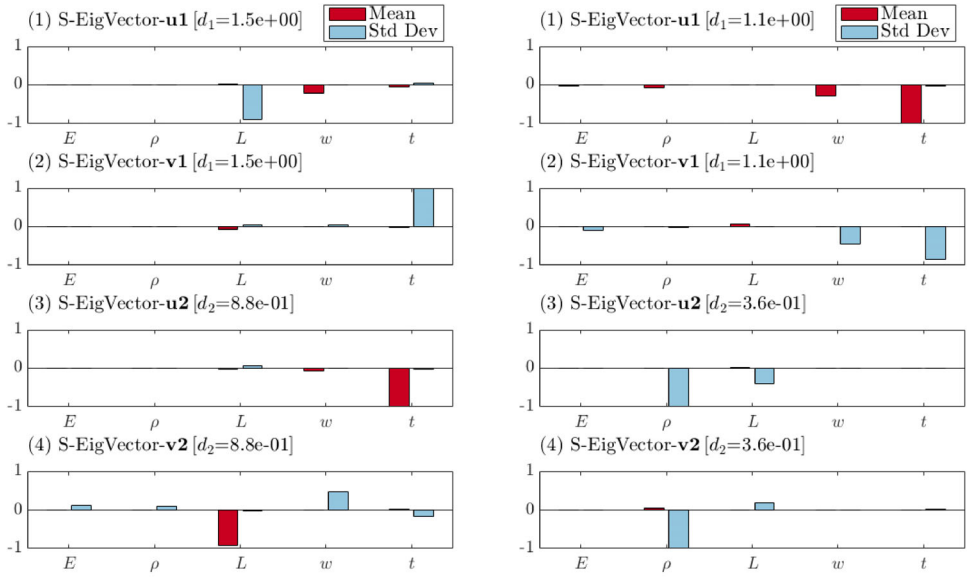


Figure 10. Case-2, same key as Figure 9.

it seems that there is a split phenomenon between the means and standard deviations of the same variables. For example, in Figure 9 (a), the first and second eigenvectors point us to the standard deviations of the variables L and t, while it is the mean values of the two variables that are important for the third and fourth eigenvectors. Similarly, in Figure 10(a) for case-2, σ_L and μ_L, the mean and standard deviation of the variable L, dominate the second and the fourth eigenvector, respectively, while the means and standard deviations of the rest are the influential parameters for the first and third eigenvectors. In other words, the dominance of the sensitivity to the mean and the standard deviation of the same variable splits into different eigenvectors, for example, σ_L dominates the first eigenvector while μ_L dominates the fourth.

This split phenomenon can be understood as a consequence of the normalization as mentioned for the motivating example in Section 1.2. The normalization compresses the relative magnitudes between the eigenvalues so that the FIM is better conditioned. This makes the ellipsoid for the relative entropy (see Section 2.2) closer to a sphere. The orthogonality of the principal axes could then result in a split between the mean and standard deviation parameters of the same variable as their influences are similar. As a result, it is difficult to identify the most influential variables. On the contrary, the symplectic



(a) S-EigVector (L and t pairing) (b) S-EigVector (L and w pairing)

Figure 11. Symplectic eigenvectors (S-EigVector) of the FIM for Case-1, for two different pairing decisions. (a) (μ_L,μ_t) and (σ_L,σ_t) in pairs; (b) (μ_L,μ_ρ) and (σ_L,σ_ρ) in pairs. The rest of the pairs are the same as Figure 9.

decomposition enforces a symplectic structure that tends to mitigate this issue by making the sensitivity information more distinctive by pairing the parameters of the same variable.

As described in Section 2, the same FIM can also be decomposed onto a symplectic basis. The symplectic eigenvalues and eigenvectors are named as "S-EigValue" and "S-EigVector" and are shown in Figures 7, 8, 9(b), and 10(b). The dimension of the symplectic eigenvalue spectrum, 5 in this case, is half of the standard eigenvalues. The symplectic eigenvectors come in pairs, $(\mathbf{u_1}, \mathbf{v_1})$ and $(\mathbf{u_2}, \mathbf{v_2})$, as shown in Figures 9(b) and 10(b), and each pair corresponds to the same symplectic eigenvalue.

As compared to the standard eigenvectors, the split parameters are grouped together in the symplectic eigenvectors. For example, for Case-1 results in Figure 9(b), the first symplectic eigenvector pairs identify L as the influential variable, with its mean and standard deviation dominating $\mathbf{u_1}$ and $\mathbf{v_1}$, respectively. The same can be found for the variable t for the second pair of symplectic eigenvectors $(\mathbf{u_2}, \mathbf{v_2})$ in Figure 9(b). Similar conclusions can be found for the Case-2 results in Figure 10(b). The grouping of the parameters is a consequence of the symplectic structure, where parameters are regarded as two-parameter pairs, for example, (μ, σ) in this case. This is pertinent in sensitivity analysis as it makes the influential variables, or two-parameter pairs, very distinctive.

It is interesting to note that in both cases, the square of the first symplectic eigenvalue is almost the same as the product of the two standard eigenvalues that split. For example, for Case-1, $d_1^2 = 1.38^2 = 1.91$ and that is about the same as the product $\lambda_1 \times \lambda_3 = 1.94 \times 0.98 = 1.90$, which corresponds to the first and fourth eigenvectors that are dominated by σ_L and μ_L, respectively.

This is a consequence of the conservation of the total sensitivity volume for symplectic decomposition as discussed in Section 2, where when two of the standard eigenvectors splits, the product of their eigenvalues tends to be conserved in the corresponding symplectic decomposition. This also occurs for decision-oriented pairings to be presented in Figure 11. For Case-1, the first and second eigenvectors are dominated by σ_L and σ_t. When these two parameters are paired together, as to be seen in Figure 11(a), $d_1^2 = 2.55$ is very similar to the product $\lambda_1 \times \lambda_2 = 2.5$ in Figure 9(a). Although in this simple example, the parameter split found from the standard eigenvalue analysis can be easily identified, it will get more difficult with a larger number of parameters. On the contrary, the parameter pairing structure is enforced by the symplectic decomposition. In addition, contrary to the standard eigenvalue decomposition where the sensitivity information is fixed for a given FIM, the symplectic variant takes account of user inputs for the pairing decisions. As an example, two different pairing decisions are considered for the same FIM from Case-1 presented in Figure 9. The symplectic eigenvectors are shown in Figure 11, with the rows and columns of the FIM rearranged as per the pairing requirement. It should be noted that while the standard eigenvalue analysis is invariant with respect to the row/column operation, the symplectic spectra are different as shown in Figure 11.

Instead of using the means and stand deviations as natural pairs for the same variables, we consider pairing the mean and standard deviation parameters for two different variables. In Figure 11(a), we pair L and t, that is, (μ_L, μ_t) and (σ_L, σ_t) in pairs, while in Figure 11(b), we pair L and ρ, that is, (μ_L, μ_ρ) and (σ_L, σ_ρ) in pairs.

It is noted in passing that although mainly for demonstrating purposes, these pairing decisions can arise in practice where the actions to reduce the uncertainties of two independent variables can impact both. For example, modifying of the production line can have the same effect on the uncertainties of the length L and the thickness t, and this would prompt a decision-oriented sensitivity analysis wrt the parameter pairs.

It is clear from Figure 11 that the sensitivity to the paired parameters are grouped together as before. For example, the sensitivity to the pair (σ_L, σ_t) dominates the first group of the symplectic eigenvectors in Figure 11(a), and (σ_L, σ_ρ) are grouped together in the second symplectic eigenvector pair in Figure 11(b).

It is interesting to note that the symplectic results in Figure 11(a) are very similar to the standard eigenvector results in Figure 9(a), although the strengths of these vectors, as indicated by the amplitudes of the eigenvalues, are different. For L and ρ pairing in Figure 11(b), the dominance of L seen in Figure 9 disappears, due to its pairing with ρ which is of low importance. This demonstrates that the symplectic decomposition is decision-oriented, as even for the same FIM, it extracts different sensitivity information according to different pairing strategies.

While only an engineering design example is considered here, the benefits of the symplectic decomposition are expected for general decision problems, whenever the spectral analysis of the FIM is used for sensitivity analysis. As the additional computation cost is negligible once the FIM is obtained, the symplectic decomposition can be used in tandem with the standard eigenvalue decomposition to extract more useful sensitivity information.

5. Conclusions

A new probabilistic sensitivity metric has been proposed based on the symplectic spectral analysis of the FIM. Contrasting to the standard eigenvalue decomposition, the symplectic decomposition of the FIM naturally identifies the sensitivity information with respect to two-parameter pairs, for example, mean and standard deviation of a random input. The resulting symplectic eigenvalues of large magnitudes, and the corresponding symplectic eigenvectors of the FIM, then reveal the most sensitive two-parameter pairs.

Through an engineering design example using a simple cantilever beam, it is observed that the normalization of the FIM tends to compress the relative magnitudes between the eigenvalues. Geometrically the relative entropy ellipsoid becomes near-spherical (see Section 2.2) due to the normalization, and this can result in a split phenomenon of different distribution parameters of the same variable. It is demonstrated that the proposed symplectic decomposition can reveal the concealed sensitivity information between the parameter pairs. Contrary to the standard eigenvalue decomposition where the sensitivity information is fixed for a given FIM, the symplectic variant takes account of user inputs for the pairing decisions. As the additional computation cost is negligible once the FIM is obtained, the symplectic decomposition can thus be used in tandem with the standard eigenvalue decomposition to gain more insight into the sensitivity information, and orient toward better decision support under uncertainties.

The proposed symplectic decomposition is only applicable for parameter space of even dimensions. For distribution parameters that belong to the two-parameter family of probability distributions, such as the widely used location-scale families, there is a natural pairing of the parameters. For more general cases, a decision-oriented two-parameter re-parameterization of the FIM is necessary and that is one of the future research to be explored.

Supplementary Materials

The supplementary materials contain details of symplectic decomposition and its computation, proofs that the symplectic egienvectors provides maximization in a symplectic basis, and codes to reproduce Figure 5.

Acknowledgments

For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising. The authors are grateful to Professor Robin Langley, University of Cambridge, for comments on an early draft of this article. We thank referees for their valuable insights and suggestions that led to an improved article.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available in the GitHub repository: *https://github.com/longitude-jyang/SymplecticFisherSensitivity*

Disclosure Statement

The authors report there are no competing interests to declare.

Funding

This work has been funded by the Engineering and Physical Sciences Research Council through the award of a Programme Grant "Digital Twins for Improved Dynamic Design," grant no. EP/R006768.

ORCID

Jiannan Yang http://orcid.org/0000-0001-8323-7406

References

Arnol'd, V. I. (1989), *Mathematical Methods of Classical Mechanics*, Graduate Texts in Mathematics (2nd ed.), New York: Springer-Verlag. DOI:10.1007/978-1-4757-2063-1. [3]

Bhatia, R., and Jain, T. (2015), "On Symplectic Eigenvalues of Positive Definite Matrices," *Journal of Mathematical Physics*, 56, 112201. [4]

Borgonovo, E. (2007), "A New Uncertainty Importance Measure," *Reliability Engineering & System Safety*, 92, 771–784. [2]

Borgonovo, E., and Plischke, E. (2016), "Sensitivity Analysis: A Review of Recent Advances," *European Journal of Operational Research*, 248, 869–887. [2]

Cover, T. M., and Thomas, J. A. (2006), *Elements of Information Theory*, Hoboken, NJ: Wiley. [Online; accessed 2022-04-20]. [2,4]

Critchfield, G. C., and Willard, K. E. (1986), "Probabilistic Analysis of Decision Trees Using Monte Carlo Simulation," *Medical Decision Making*, 6, 85–92. [2]

Fisher, R. A., and Russell, E. J. (1922), "On the Mathematical Foundations of Theoretical Statistics," *Philosophical Transactions of the Royal Society of London, Series A, Containing Papers of a Mathematical or Physical Character*, 222, 309–368. [1]

Gauchy, C., Stenger, J., Sueur, R., and Iooss, B. (2022), "An Information Geometry Approach to Robustness Analysis for the Uncertainty Quantification of Computer Codes," *Technometrics*, 64, 80–91. DOI:10.1080/00401706.2021.1905072. [4]

Gosson, d. M. (2006), *Symplectic Geometry and Quantum Mechanics / Maurice de Gosson*, Operator Theory, Advances and Applications (Vol. 166), Basel: Birkhäuser. [4]

Gunawan, R., Cao, Y., Petzold, L., and Doyle, F. J. (2005), "Sensitivity Analysis of Discrete Stochastic Systems," *Biophysical Journal*, 88, 2530–2540. [1]

Jia, G., and Taflanidis, A. A. (2014), "Sample-based Evaluation of Global Probabilistic Sensitivity Measures," *Computers & Structures*, 144, 103–118. [2]

Komorowski, M., Costa, M. J., Rand, D. A., and Stumpf, M. P. H. (2011), "Sensitivity, Robustness, and Identifiability in Stochastic Chemical Kinetics Models," *Proceedings of the National Academy of Sciences*, 108, 8645–8650. [1]

Lee, E. D., Katz, D. M., Bommarito, M. J., and Ginsparg, P. H. (2020), "Sensitivity of Collective Outcomes Identifies Pivotal Components," *Journal of The Royal Society Interface*, 17, 20190873. [1]

Lehmann, E. L., and Casella, G. (1998), *Theory of Point Estimation*, Springer Texts in Statistics (2nd ed.), New York: Springer. [2]

Lemaître, P., Sergienko, E., Arnaud, A., Bousquet, N., Gamboa, F., and Iooss, B. (2015), "Density Modification-based Reliability Sensitivity Analysis," *Journal of Statistical Computation and Simulation*, 85, 1200–1223. [2]

Majda, A. J., and Gershgorin, B. (2010), "Quantifying Uncertainty in Climate Change Science through Empirical Information Theory," *Proceedings of the National Academy of Sciences*, 107, 14958–14963. [1]

Nicacio, F. (2021), "Williamson Theorem in Classical, Quantum, and Statistical Physics," *American Journal of Physics*, 89, 1139–1151. [3,4]

Oakley, J. E. (2009), "Decision-Theoretic Sensitivity Analysis for Complex Computer Models," *Technometrics*, 51, 121–129. DOI:10.1198/TECH.2009.0014. [1]

Oakley, J. E., and O'Hagan, A. (2004), "Probabilistic Sensitivity Analysis of Complex Models: A Bayesian Approach," *Journal of the Royal Statistical Society*, Series B, 66, 751–769. *https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9868.2004.05304.x*. [1,2,4,5,6,7]

Pantazis, Y., Katsoulakis, M. A., and Vlachos, D. G. (2013), "Parametric Sensitivity Analysis for Biochemical Reaction Networks based on Pathwise Information Theory," *BMC Bioinformatics*, 14, 311. [3]

Pereira, J. L., Banchi, L., and Pirandola, S. (2021), "Symplectic Decomposition from Submatrix Determinants," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477, 20210513. [3]

Rubinstein, R. Y., and Kroese, D. P. (2016), *Simulation and the Monte Carlo Method* (3rd ed.), New York: Wiley. [5]

Saltelli, A., ed. (2008), *Global Sensitivity Analysis: The Primer*, Chichester, England; Hoboken, NJ: Wiley. [2,5]

Sobol', I., and Kucherenko, S. (2009), "Derivative based Global Sensitivity Measures and their Link with Global Sensitivity Indices," *Mathematics and Computers in Simulation*, 79, 3009–3017. [2]

Williamson, J. (1936), "On the Algebraic Problem Concerning the Normal Forms of Linear Dynamical Systems," *American Journal of Mathematics*, 58, 141–163. [3]

Yang, J. (2022), "An Information Upper Bound for Probability Sensitivity," arXiv:2206.02274 [cs, math, stat]. [4]

——— (2023), "A General Framework for Probabilistic Sensitivity Analysis with Respect to Distribution Parameters," *Probabilistic Engineering Mechanics*, 72, 103433. [2,4,5,8]

Yang, J., Langley, R. S., and Andrade, L. (2022), "Digital Twins for Design in the Presence of Uncertainties," *Mechanical Systems and Signal Processing*, 179, 109338. [1,2,3,4,5,8]