

An Unified Parametric-Nonparametric Uncertainty Quantification Approach for Linear Dynamical Systems

S ADHIKARI

School of Engineering, University of Wales Swansea, Swansea, U.K.

Email: S.Adhikari@swansea.ac.uk

URL: <http://engweb.swan.ac.uk/~adhikaris>

Outline of the presentation

- Uncertainty in structural dynamics
- Critical review of current UQ approaches
- Random matrix models
- Derivation of noncentral Wishart distribution
- Numerical implementations and example
- Conclusions & discussions

Overview of predictive approaches

There are five key steps:

- Physics (mechanics) model building
- Uncertainty Quantification (UQ)
- Uncertainty Propagation (UP)
- Model Verification & Validation (V & V)
- Prediction

Tools are available for each of these steps. My focus in this talk is on UQ in linear dynamical systems.

Complex aerospace models



Possible uncertain subsystems of an aircraft

Why uncertainty?

Different sources of uncertainties in the **modeling** and **parameters** of dynamic systems may be attributed, but not limited, to the following factors:

- **Mathematical models:** equations (linear, non-linear), geometry, damping model (viscous, non-viscous, fractional derivative), boundary conditions/initial conditions, input forces;
- **Model parameters:** Young's modulus, mass density, Poisson's ratio, damping model parameters (damping coefficient, relaxation modulus, fractional derivative order)

Why uncertainty?

- **Numerical algorithms:** weak formulations, discretisation of displacement fields (in finite element method), discretisation of stochastic fields (in stochastic finite element method), approximate solution algorithms, truncation and roundoff errors, tolerances in the optimization and iterative methods, artificial intelligent (AI) method (choice of neural networks)
- **Measurements:** noise, resolution (number of sensors and actuators), experimental hardware, excitation method (nature of shakers and hammers), excitation and measurement point, data processing (amplification, number of data points, FFT), calibration

Structural dynamics

The equation of motion:

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{f}(t) \quad (1)$$

- Due to the presence of uncertainty \mathbf{M} , \mathbf{C} and \mathbf{K} become random matrices.
- The main objectives in the ‘forward problem’ are:
 - to quantify uncertainties in the system matrices
 - to predict the variability in the response vector \mathbf{x}

Current UQ approaches

Two different approaches are currently available

- **Parametric approaches**: Such as the **Stochastic Finite Element Method (SFEM)**:
 - aim to characterize **aleatoric** uncertainty
 - assumes that stochastic fields describing parametric uncertainties are known in details
 - suitable for low-frequency dynamic applications

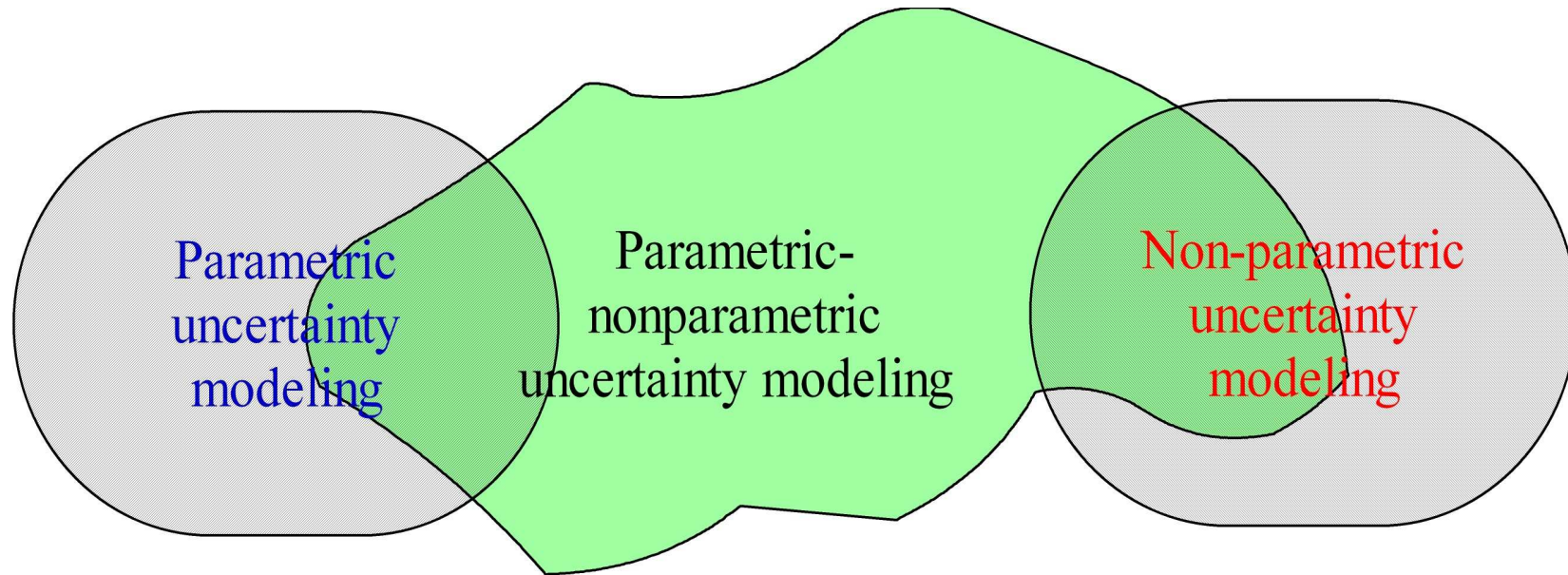
Current UQ approaches

- Nonparametric approaches : Such as the Statistical Energy Analysis (SEA) and Wishart random matrix theory:
 - aim to characterize episematic uncertainty
 - does not consider parametric uncertainties in details
 - suitable for high-frequency dynamic applications

Limitations of current UQ approaches

- Although we have mentioned and made differences between the two different types of uncertainties, in practical problems it is in general very difficult, if not impossible, to distinguish them.
- Recently reported experimental studies by our group on one hundred nominally identical beams and plates emphasize this fact.
- For credible numerical models of complex dynamical systems, we need to quantify and model both types of uncertainties **simultaneously**.
- **A hybrid approach is required.**

Overview of proposed approach



Stochastic FE

Random Matrix Theory
(central Wishart distribution)

Schematic representation of the proposed parametric-nonparametric uncertainty modeling in structural dynamics.

Proposed unified approach

- **The objective**: To develop a **hybrid approach** which takes both parametric and nonparametric uncertainties into account.
- **The rationale**: No matter what the nature of uncertainty is (parametric/nonparametric or both), at the end it will result in random M , C and K matrices.
- **The methodology**: Derive the matrix variate probability density functions of M , C and K based on parametric information (e.g. mean and covariance of the elements) and overall physically realistic mathematical constraints (such as the symmetry and positive definiteness).

Matrix variate distributions

- The probability density function of a random matrix can be defined in a manner similar to that of a random variable.
- If \mathbf{A} is an $n \times m$ real random matrix, the matrix variate probability density function of $\mathbf{A} \in \mathbb{R}_{n,m}$, denoted as $p_{\mathbf{A}}(\mathbf{A})$, is a mapping from the space of $n \times m$ real matrices to the real line, i.e., $p_{\mathbf{A}}(\mathbf{A}) : \mathbb{R}_{n,m} \rightarrow \mathbb{R}$.

Gaussian random matrix

The random matrix $\mathbf{X} \in \mathbb{R}_{n,p}$ is said to have a matrix variate Gaussian distribution with mean matrix $\mathbf{M} \in \mathbb{R}_{n,p}$ and covariance matrix $\Sigma \otimes \Psi$, where $\Sigma \in \mathbb{R}_n^+$ and $\Psi \in \mathbb{R}_p^+$ provided the pdf of \mathbf{X} is given by

$$p_{\mathbf{X}}(\mathbf{X}) = (2\pi)^{-np/2} |\Sigma|^{-p/2} |\Psi|^{-n/2} \operatorname{etr} \left\{ -\frac{1}{2} \Sigma^{-1} (\mathbf{X} - \mathbf{M}) \Psi^{-1} (\mathbf{X} - \mathbf{M})^T \right\} \quad (2)$$

This distribution is usually denoted as $\mathbf{X} \sim N_{n,p}(\mathbf{M}, \Sigma \otimes \Psi)$.

Central Wishart matrix

A $n \times n$ symmetric positive definite random matrix \mathbf{S} is said to have a Wishart distribution with parameters $p \geq n$ and $\Sigma \in \mathbb{R}_n^+$, if its pdf is given by

$$p_{\mathbf{S}}(\mathbf{S}) = \left\{ 2^{\frac{1}{2}np} \Gamma_n \left(\frac{1}{2}p \right) |\Sigma|^{\frac{1}{2}p} \right\}^{-1} |\mathbf{S}|^{\frac{1}{2}(p-n-1)} \text{etr} \left\{ -\frac{1}{2} \Sigma^{-1} \mathbf{S} \right\} \quad (3)$$

This distribution is usually denoted as $\mathbf{S} \sim W_n(p, \Sigma)$.

Note: This distribution is used in current nonparametric UQ methods.

Noncentral Wishart matrix

A $n \times n$ symmetric positive definite random matrix \mathbf{S} is said to have a noncentral Wishart distribution with parameters $p \geq n$, $\Sigma \in \mathbb{R}_n^+$ and $\Theta \in \mathbb{R}_n^+$, if its pdf is given by

$$p_{\mathbf{S}}(\mathbf{S}) = \left\{ 2^{\frac{1}{2}np} \Gamma_n \left(\frac{1}{2}p \right) |\Sigma|^{\frac{1}{2}p} \right\}^{-1} \text{etr} \left\{ -\frac{1}{2}\Theta \right\} \text{etr} \left\{ -\frac{1}{2}\Sigma^{-1}\mathbf{S} \right\} \\ |\mathbf{S}|^{\frac{1}{2}(p-n-1)} {}_0F_1(p/2, \Theta\Sigma^{-1}\mathbf{S}/4). \quad (4)$$

where ${}_0F_1$ the hypergeometric function (Bessel function) of a matrix argument. This distribution is usually denoted as $\mathbf{S} \sim W_n(p, \Sigma, \Theta)$. Note that if the noncentrality parameter Θ is a **null matrix**, then it reduces to the central Wishart distribution.



Distribution of the system matrices

The distribution of the random system matrices \mathbf{M} , \mathbf{C} and \mathbf{K} should be such that they are

- symmetric
- positive-definite, and
- the moments (at least first two) of the inverse of the dynamic stiffness matrix

$$\mathbf{D}(\omega) = -\omega^2\mathbf{M} + i\omega\mathbf{C} + \mathbf{K} \text{ should exist } \forall \omega$$

Current nonparametric approach

- Suppose $\mathbf{G} \equiv \{\mathbf{M}, \mathbf{C}, \mathbf{K}\}$
- $\mathbf{G} \sim W_n(p, \Sigma)$ where $p = n + 1 + \theta$, $\Sigma = \overline{\mathbf{G}} / \sqrt{\theta(n + 1 + \theta)}$
and $\theta = \frac{1}{\delta_G^2} \left\{ 1 + \frac{\{\text{Trace}(\overline{\mathbf{G}})\}^2}{\text{Trace}(\overline{\mathbf{G}}^2)} \right\} - (n + 1)$
- $\delta_G^2 = \frac{\mathbb{E}[\|\mathbf{G} - \mathbb{E}[\mathbf{G}]\|_{\text{F}}^2]}{\|\mathbb{E}[\mathbf{G}]\|_{\text{F}}^2} = \frac{\text{Trace}(\text{cov}(\text{vec}(\mathbf{G})))}{\text{Trace}(\overline{\mathbf{G}}^2)}$ (normalized std) .
- **The main limitation:** $\text{cov}(G_{ij}, G_{kl}) = \frac{1}{\theta} (\overline{G}_{ik}\overline{G}_{jl} + \overline{G}_{il}\overline{G}_{jk})$
- Only **one** parameter controls the uncertainty

Current nonparametric approach

- The covariance matrix of G can have $n(n + 1) \times (n(n + 1) + 2)/8$ number of independent parameters.
- Current nonparametric approach, **only offers a single parameter** to quantify uncertainty which can potentially be expressed by $n(n + 1)(n(n + 1) + 2)/8$ number of independent parameters - **a gross oversimplification**.
- To account for parametric uncertainties, we need a matrix variate distribution which not only satisfy the mathematical constraints, but also **must offer more parameters** to fit the 'known' covariance tensor of G .

Matrix factorization approach

- Because \mathbf{G} is a symmetric and positive-definite random matrix, it can be always factorized as

$$\mathbf{G} = \mathbf{X}\mathbf{X}^T \quad (5)$$

where $\mathbf{X} \in \mathbb{R}^{n \times p}$, $p \geq n$ is in general a rectangular matrix.

- Extending the standard maximum entropy argument to the matrix case we can say that the pdf of \mathbf{X} is given by the matrix variate Gaussian distribution, that is,
$$\mathbf{X} \sim N_{n,p}(\mathcal{M}, \Sigma \otimes \mathbf{I}_p).$$
- This shows that \mathbf{G} has non central Wishart distribution.

The main result

Theorem 1. *The unified parametric-nonparametric probability density function a random system matrix $\mathbf{G} \equiv \{\mathbf{M}, \mathbf{C}, \mathbf{K}\}$ follows the noncentral Wishart distribution, that is $\mathbf{G} \sim W_n(p, \mathbf{\Sigma}, \mathbf{\Theta})$ where $p > n$ is a real scalar, $\mathbf{\Sigma}$ and $\mathbf{\Theta}$ are symmetric positive-definite $n \times n$ real matrices.*

Noncentral distribution

- If the noncentrality parameter Θ is a null matrix, the unified distribution reduces to the nonparametric distribution (central Wishart distribution).
- The unified distribution derived here is therefore further generalization of the nonparametric distribution.
- The additional $n(n + 1)/2$ parameters provided by the matrix $\Theta \in \mathbb{R}_n^+$ allow to model parametric uncertainty which is not available within the scope of the nonparametric distribution.

Parameter estimation

- We match the mean and covariance of the distribution of \mathbf{G} with 'measured/known' quantities.

$$E[\mathbf{G}] = p\mathbf{\Sigma} + \mathbf{\Omega},$$

$$\text{cov}(\text{vec}(\mathbf{G})) = (\mathbf{I}_{n^2} + \mathbf{K}_{nn}) (p\mathbf{\Sigma} \otimes \mathbf{\Sigma} + \mathbf{\Omega} \otimes \mathbf{\Sigma} + \mathbf{\Sigma} \otimes \mathbf{\Omega}).$$

- Mean is satisfied exactly while the covariance is satisfied in least-square sense.
- Suppose $\bar{\mathbf{G}} \in \mathbb{R}_n^+$, the mean matrix and $\mathbf{C}_G = \text{cov}(\text{vec}(\mathbf{G})) \in \mathbb{R}_{n^2}^+$, the covariance matrix, are known.

Parameter estimation

- Obtain the normalized standard deviation δ_G of \mathbf{G} :

$$\delta_G^2 = \frac{\mathbb{E}[\|\mathbf{G} - \mathbb{E}[\mathbf{G}]\|_F^2]}{\|\mathbb{E}[\mathbf{G}]\|_F^2} = \frac{\text{Trace}(\mathbf{C}_G)}{\text{Trace}(\overline{\mathbf{G}}^2)}$$

- $$p = \frac{1}{\delta_G^2} \frac{\text{Trace}(\overline{\mathbf{G}}^2) + \{\text{Trace}(\overline{\mathbf{G}})\}^2}{\text{Trace}(\overline{\mathbf{G}}^2)}$$

- Form the matrix $\mathcal{A} = \overline{\mathbf{G}} \otimes \overline{\mathbf{G}} - p\mathbf{C}_G/2 \in \mathbb{R}^{n^2 \times n^2}$ and obtain $\Omega \in \mathbb{R}^{n \times n}$ by least-square minimization of the Frobenius norm $\|\mathcal{A} - \Omega \otimes \Omega\|_F$.

- Calculate $\Sigma = (\overline{\mathbf{G}} - \Omega) / p$ and $\Theta = \Sigma^{-1}\Omega$.

Numerical recipe

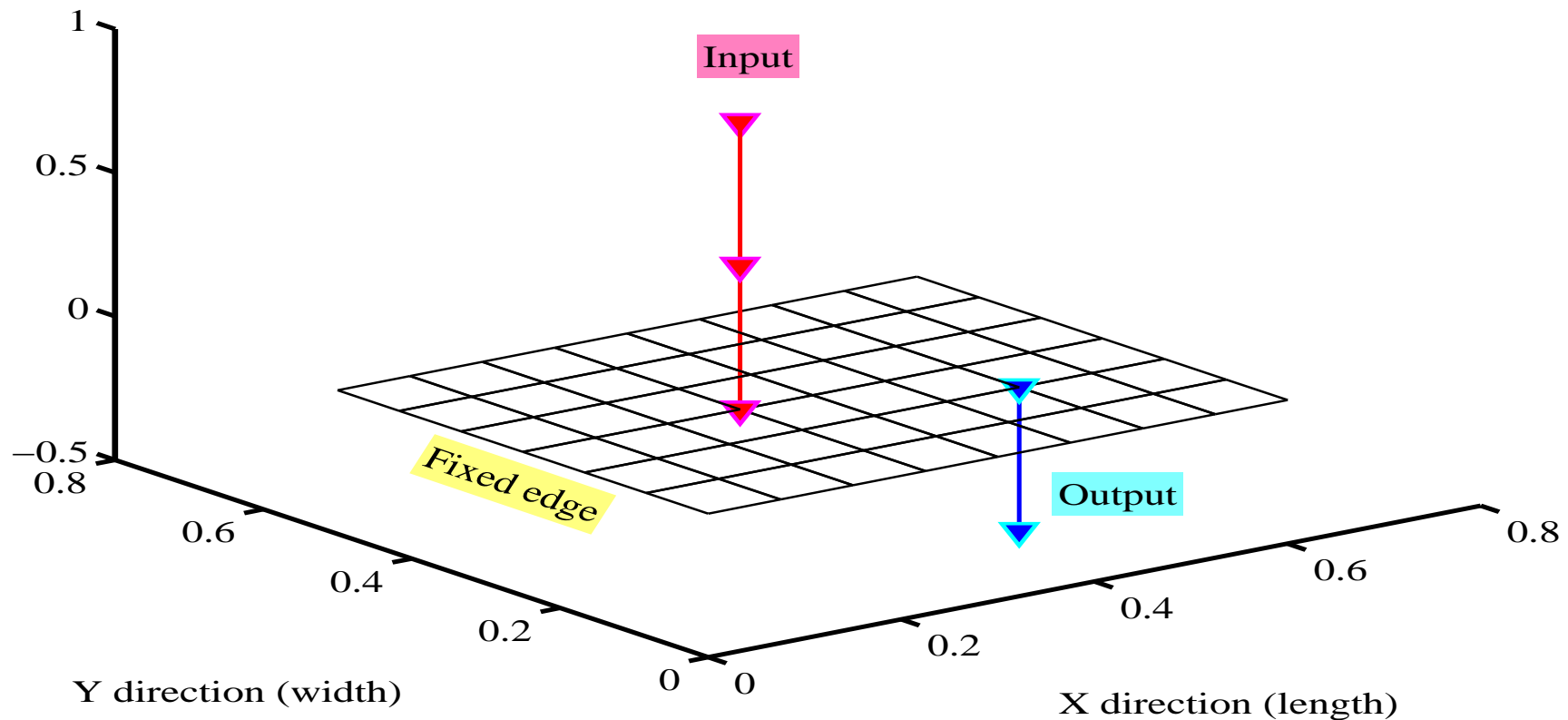
- Obtain the distribution parameters $p \in \mathbb{R}$, $\Sigma \in \mathbb{R}_n^+$ and $\Omega \in \mathbb{R}_n^+$ from \bar{G} and C_G
- Perform the Cholesky factorizations of the positive definite matrices $\Sigma \in \mathbb{R}_n^+$ and $\Omega \in \mathbb{R}_n^+$ as $\Sigma = \mathbf{D}\mathbf{D}^T$, $\mathbf{D} \in \mathbb{R}^{n \times n}$ and $\Omega = \widehat{\mathcal{M}}\widehat{\mathcal{M}}^T$, $\widehat{\mathcal{M}} \in \mathbb{R}^{n \times n}$.
- Calculate the $n \times n$ square matrix $\widetilde{\mathcal{M}} = \mathbf{D}^{-1}\widehat{\mathcal{M}}$
- Construct the $n \times p$ rectangular mean matrix $\mathcal{M} = [\widetilde{\mathcal{M}}, \mathbf{O}_{n,n-p}] \in \mathbb{R}^{n \times p}$.

Numerical recipe

- Obtain the matrix $\mathbf{Y} \in \mathbb{R}^{n \times p}$ containing uncorrelated Gaussian random numbers with mean \mathcal{M} and unit standard deviation.
- Generate the samples of a system matrix as $\mathbf{G} = \mathbf{D}\mathbf{Y}\mathbf{Y}^T\mathbf{D}^T \in \mathbb{R}_n^+$.
- In MATLAB[®], the following four lines of code will generate the samples of the system matrices:

```
D=[chol(Sigma)]'; Mhat=[chol(Omega)]';  
Mtilde=D\Mhat;  
Y=[Mtilde zeros(n,p-n)] + randn(n,p);  
G=D*Y*Y'*D';
```

Example 1: A cantilever plate



A steel cantilever plate: 8×6 elements, 168 degrees-of-freedom; $\bar{E} = 200 \times 10^9 \text{ N/m}^2$, $\bar{\mu} = 0.3$,

$\bar{\rho} = 7860 \text{ kg/m}^3$, $\bar{t} = 3.0 \text{ mm}$, $L_x = 0.6 \text{ m}$, $L_y = 0.5 \text{ m}$, 2% modal damping factor.



Stochastic properties

The Young's modulus, Poissons ratio, mass density and thickness are random fields of the form

$$E(\mathbf{x}) = \bar{E} (1 + \epsilon_E f_1(\mathbf{x})) \quad (6)$$

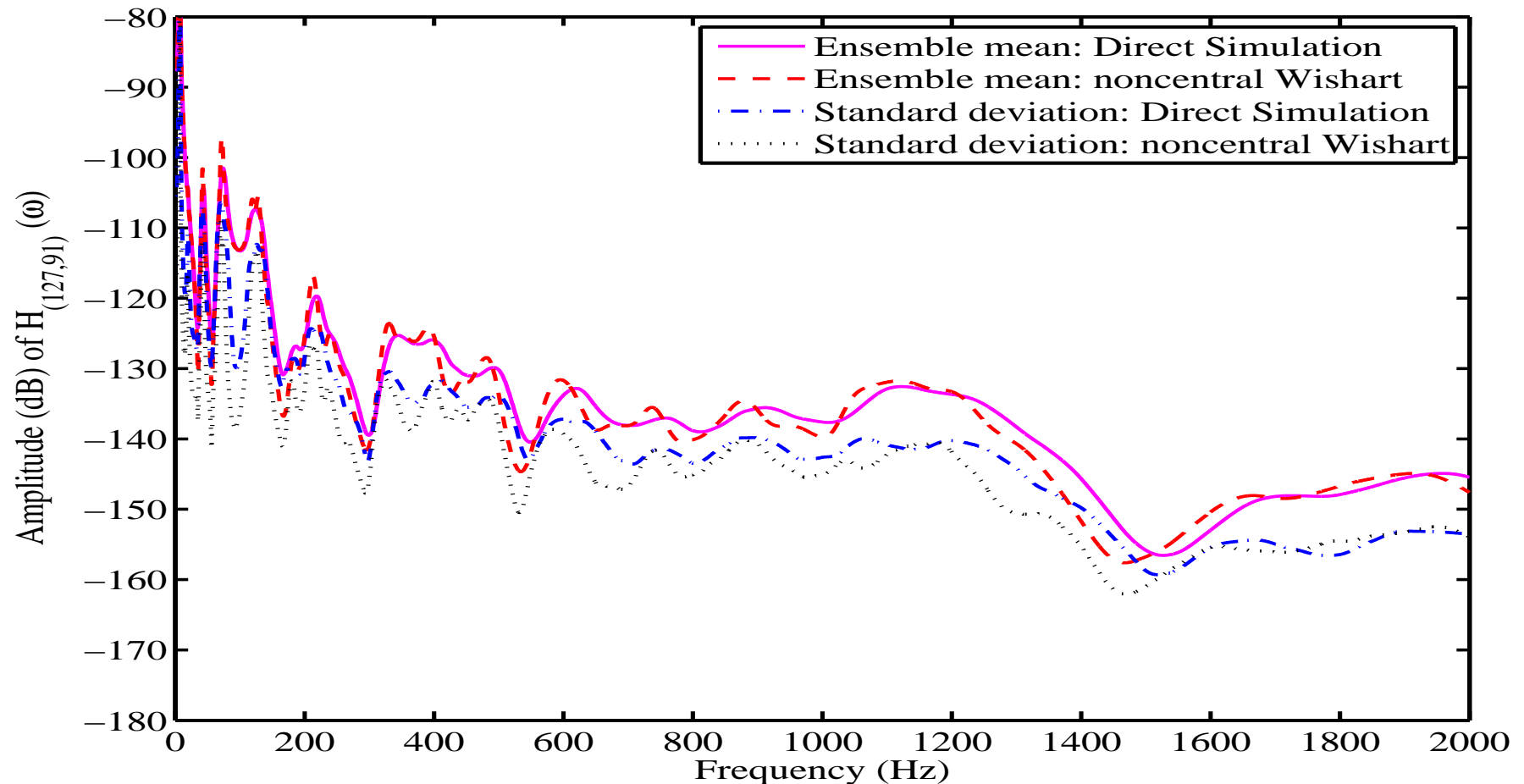
$$\mu(\mathbf{x}) = \bar{\mu} (1 + \epsilon_\mu f_2(\mathbf{x})) \quad (7)$$

$$\rho(\mathbf{x}) = \bar{\rho} (1 + \epsilon_\rho f_3(\mathbf{x})) \quad (8)$$

$$\text{and } t(\mathbf{x}) = \bar{t} (1 + \epsilon_t f_4(\mathbf{x})) \quad (9)$$

- The strength parameters: $\epsilon_E = 0.15$, $\epsilon_\mu = 0.10$, $\epsilon_\rho = 0.15$ and $\epsilon_t = 0.15$.
- The random fields $f_i(\mathbf{x})$, $i = 1, \dots, 4$ are delta-correlated homogenous Gaussian random fields.

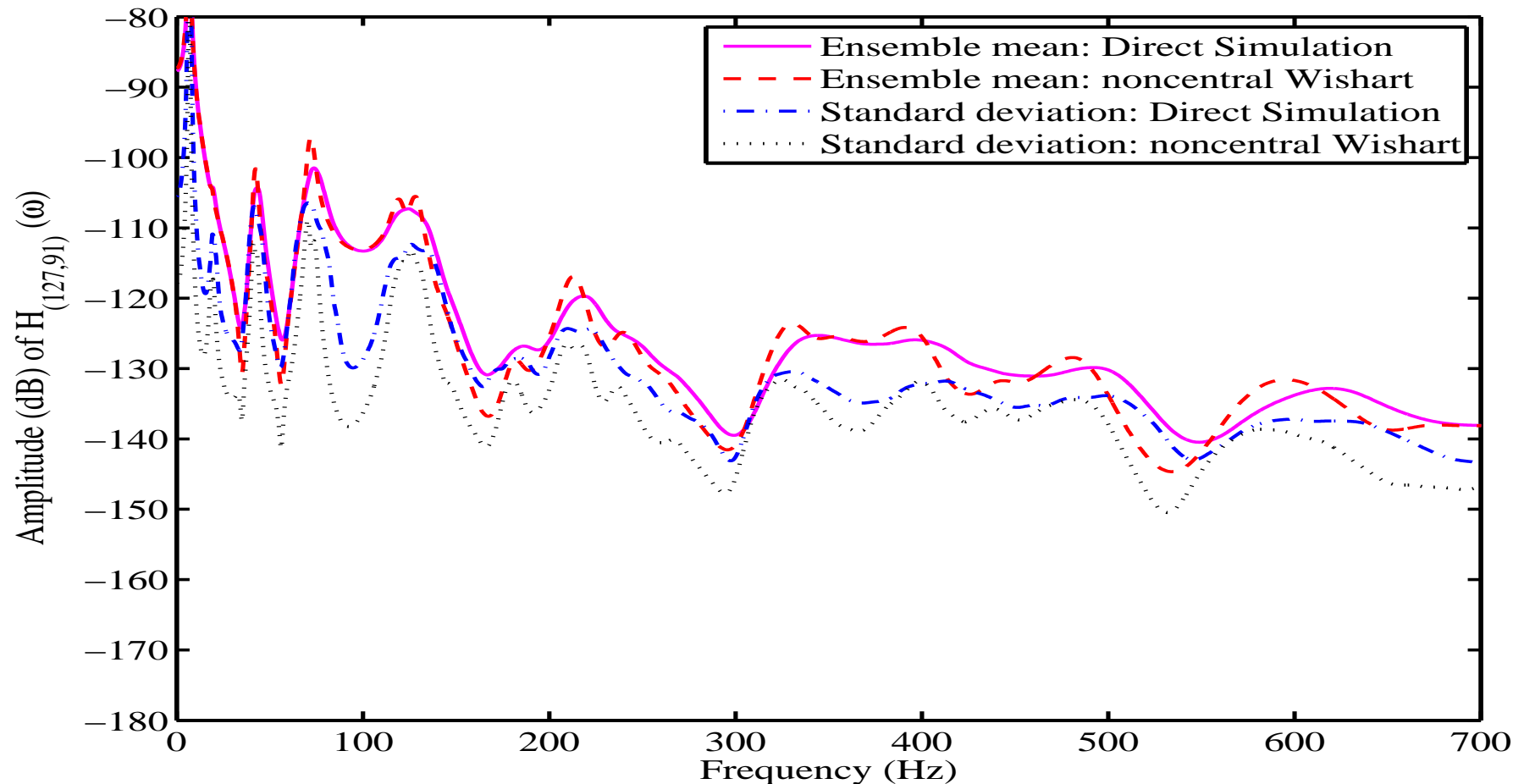
Comparison of cross-FRF



Comparison of the mean and standard deviation of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

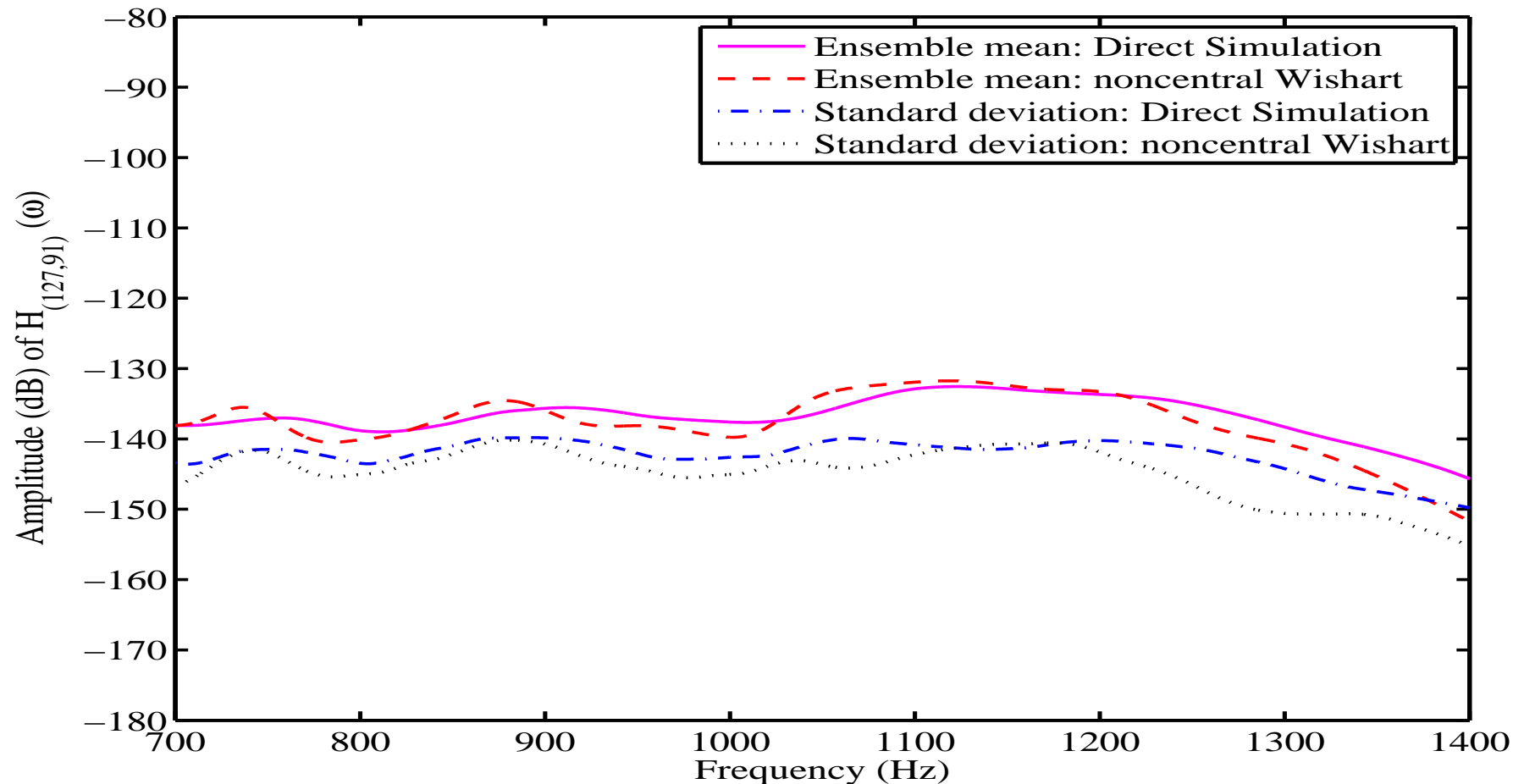
Comparison of cross-FRF: Low Freq



Comparison of the mean and standard deviation of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

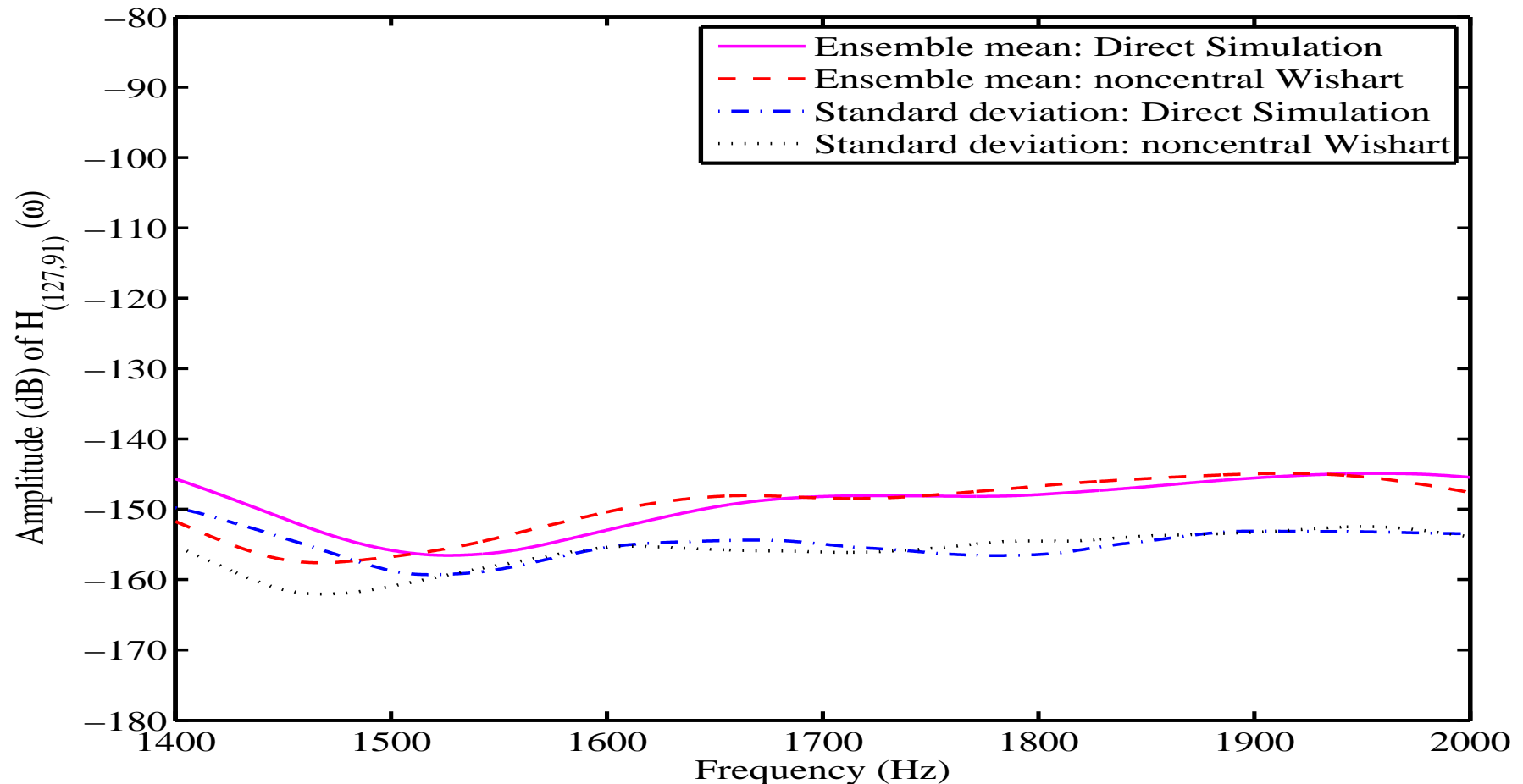
Comparison of cross-FRF: Mid Freq



Comparison of the mean and standard deviation of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

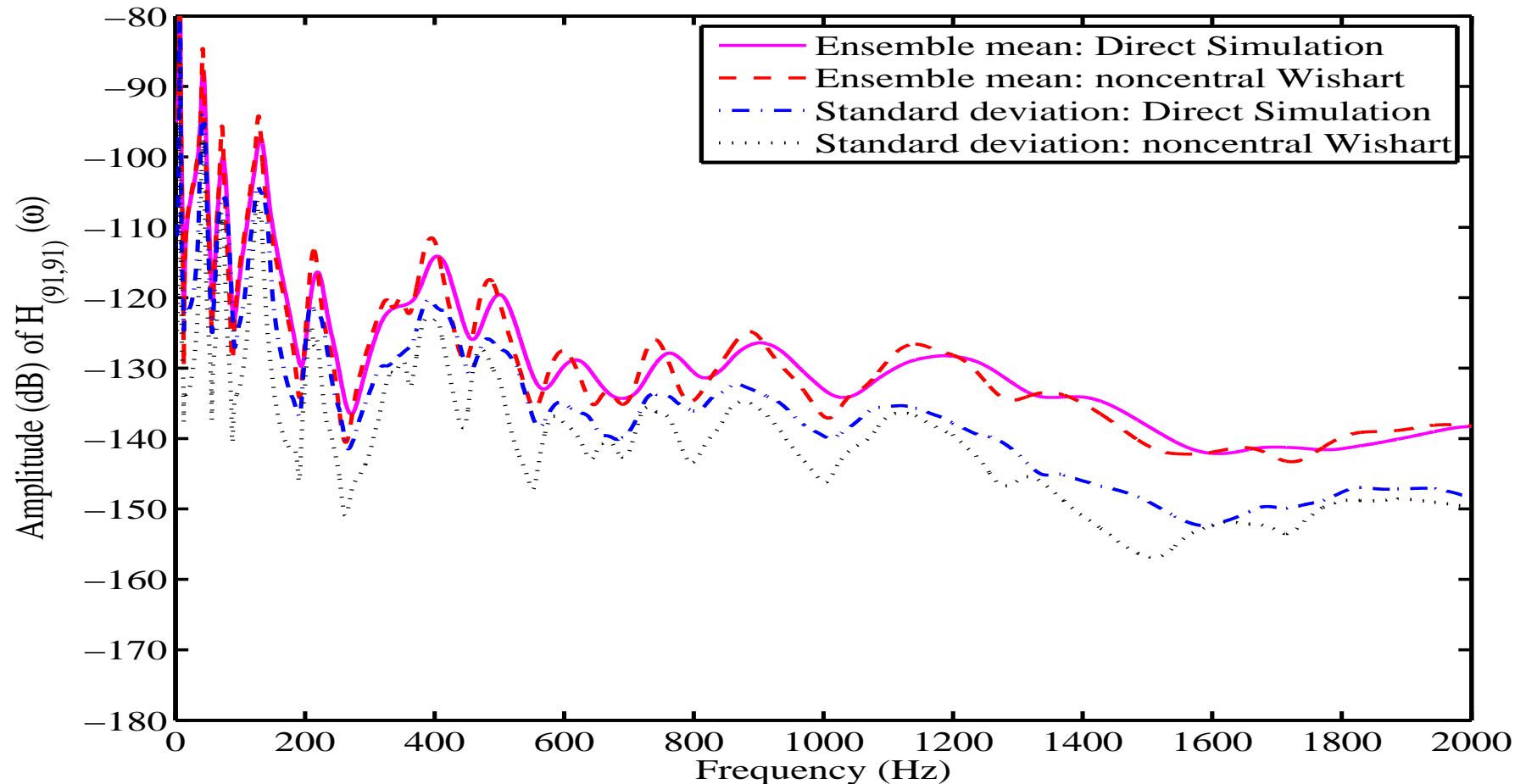
Comparison of cross-FRF: High Freq



Comparison of the mean and standard deviation of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

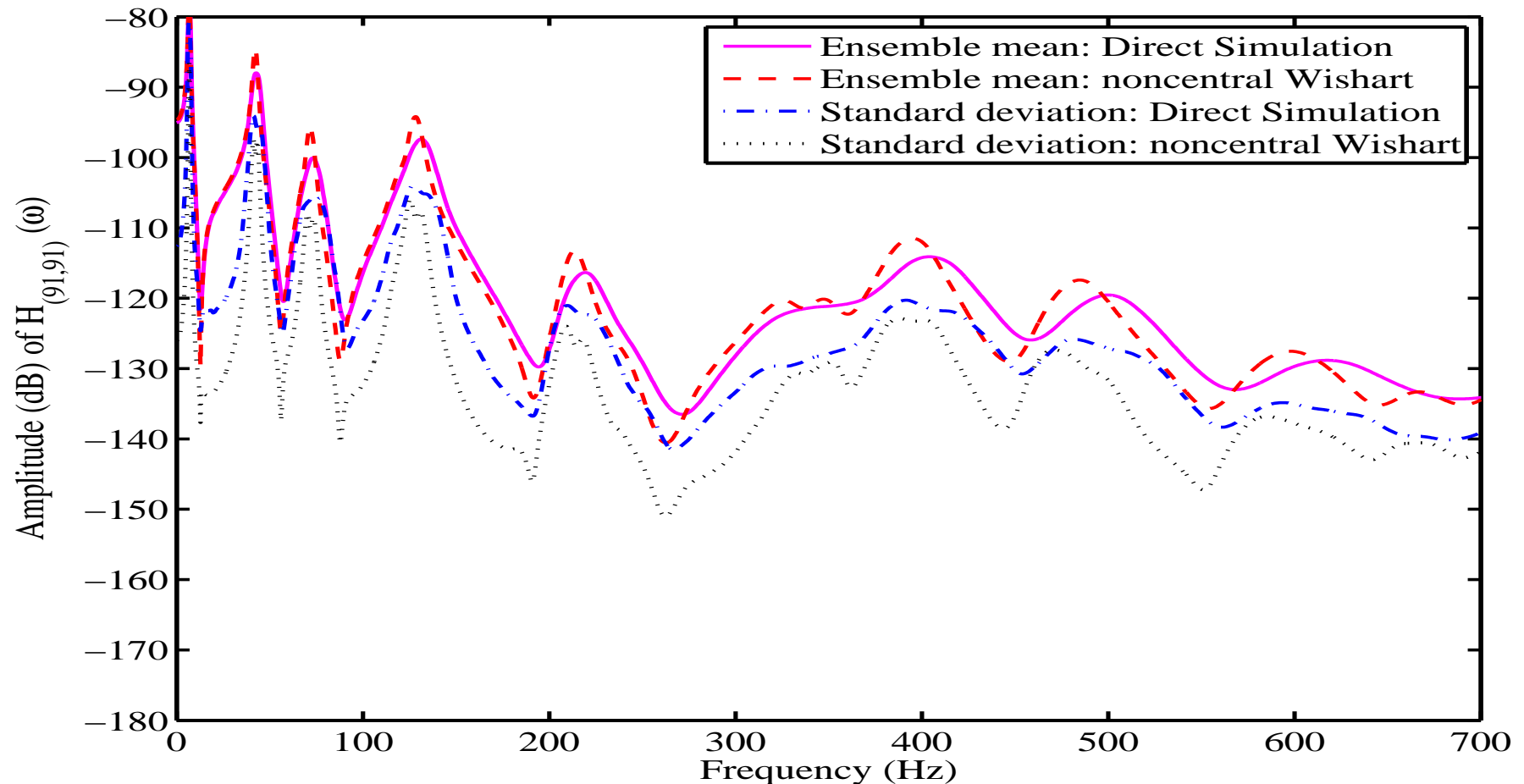
Comparison of driving-point-FRF



Comparison of the mean and standard deviation of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

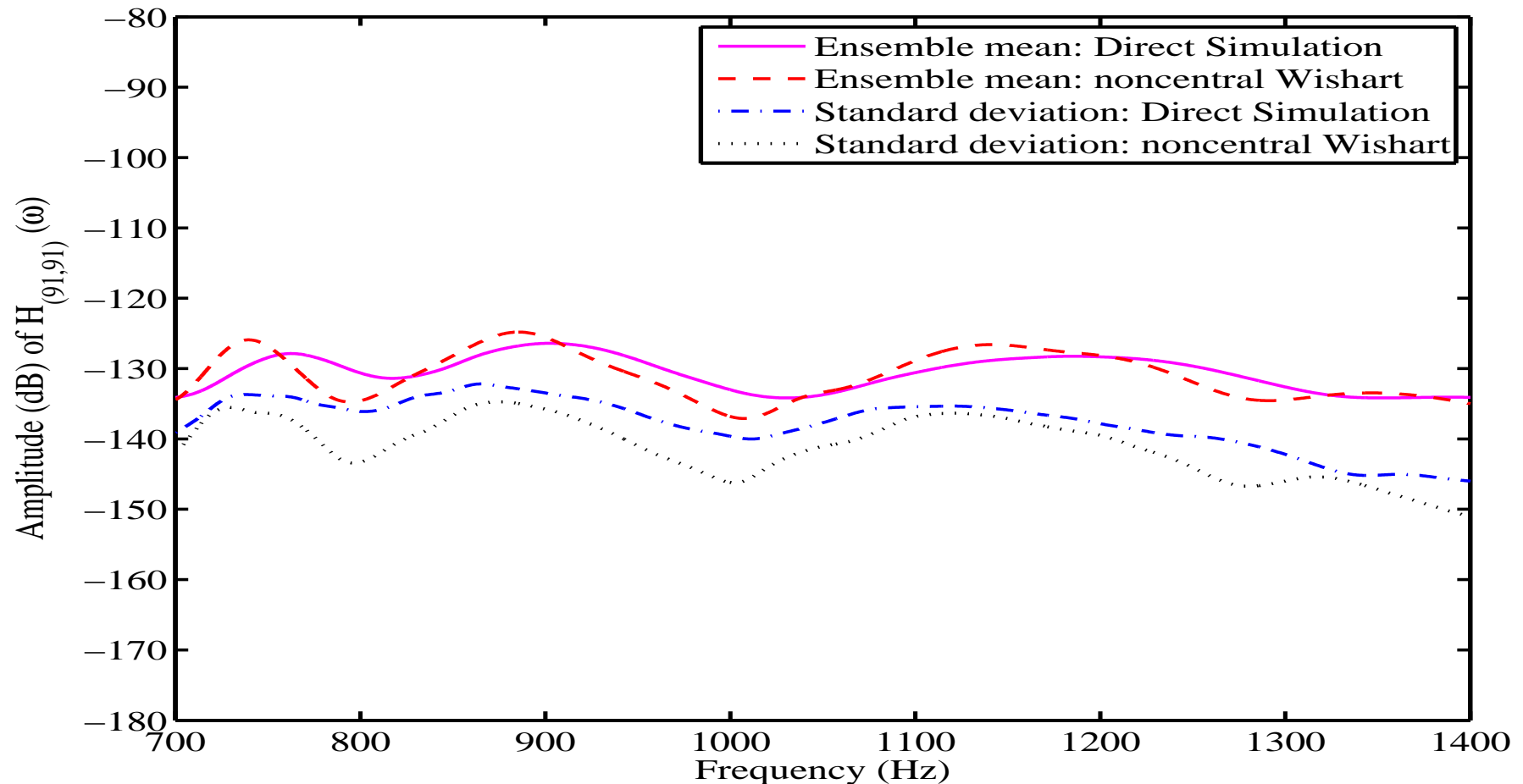
Comparison of driving-point-FRF: Low Freq



Comparison of the mean and standard deviation of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

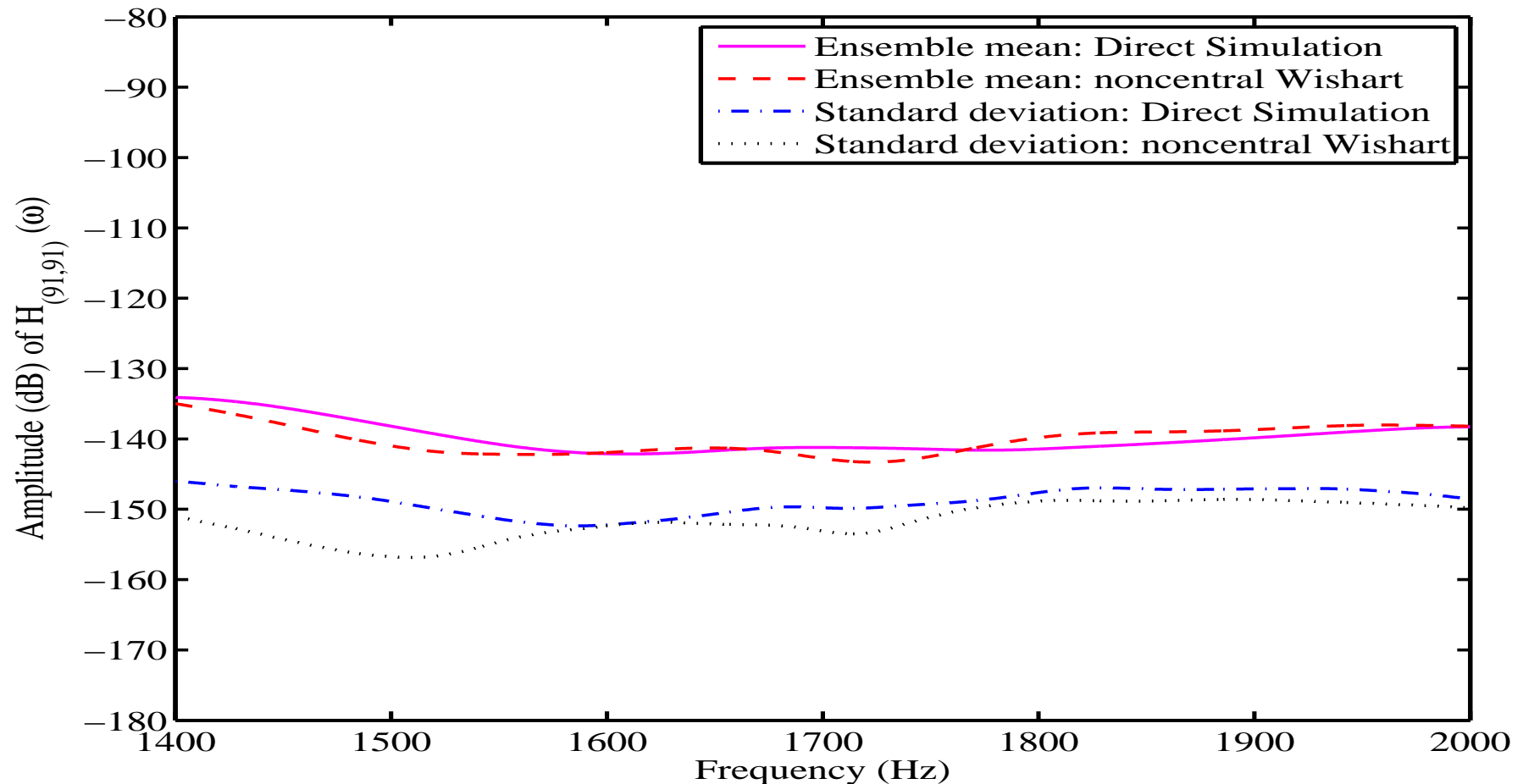
Comparison of driving-point-FRF: Mid Freq



Comparison of the mean and standard deviation of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

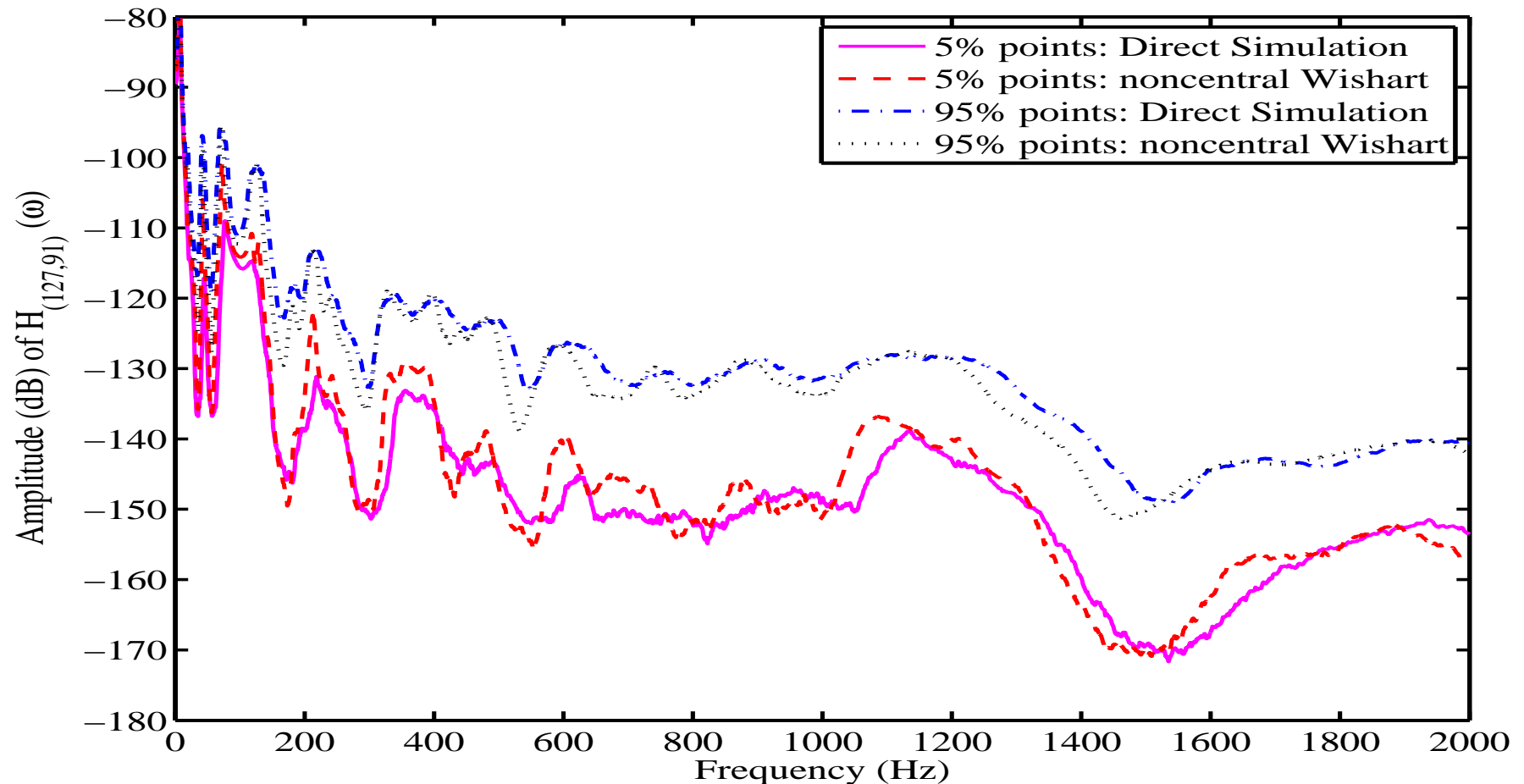
Comparison of driving-point-FRF: High Freq



Comparison of the mean and standard deviation of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

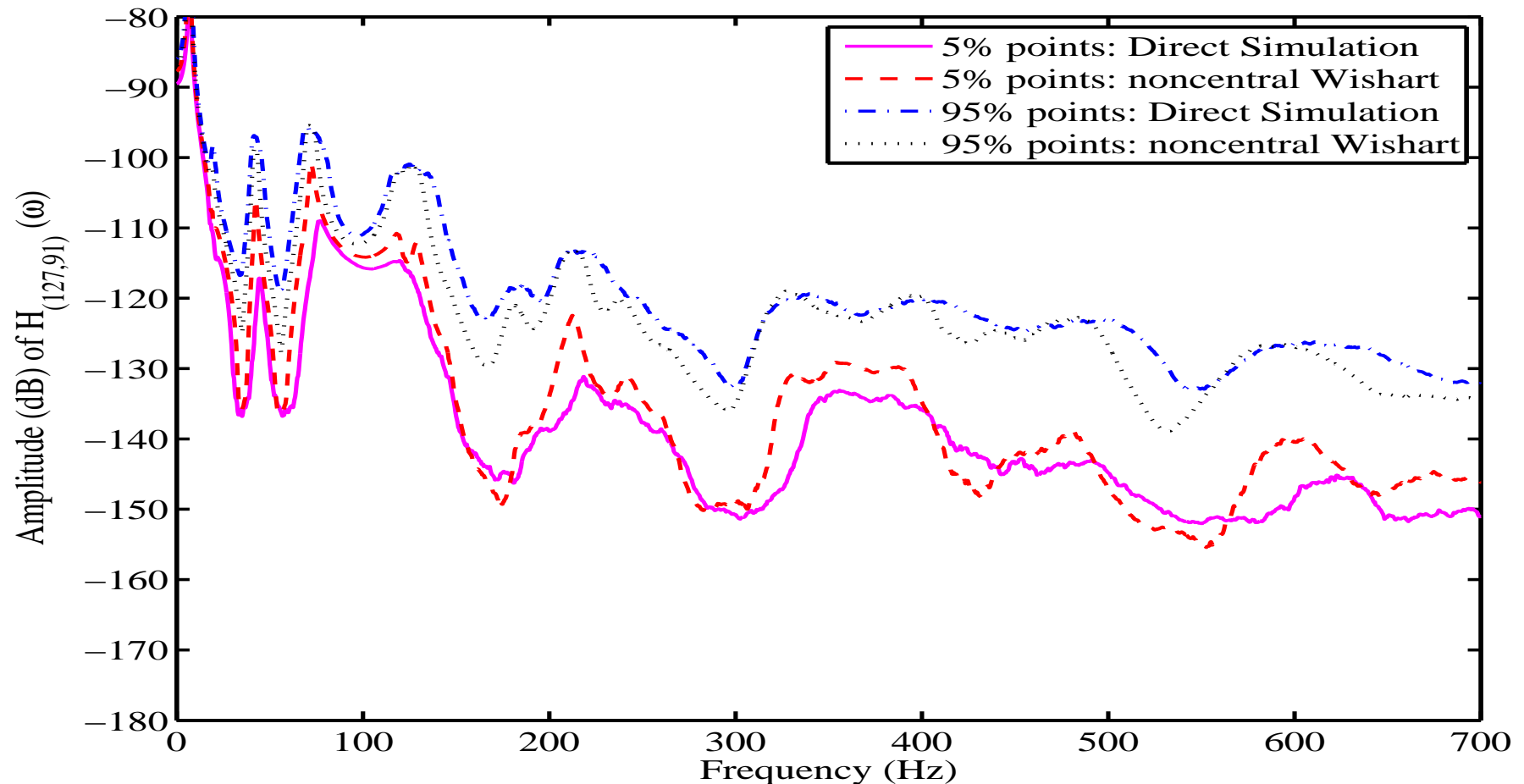
Comparison of cross-FRF



Comparison of the 5% and 95% probability points of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

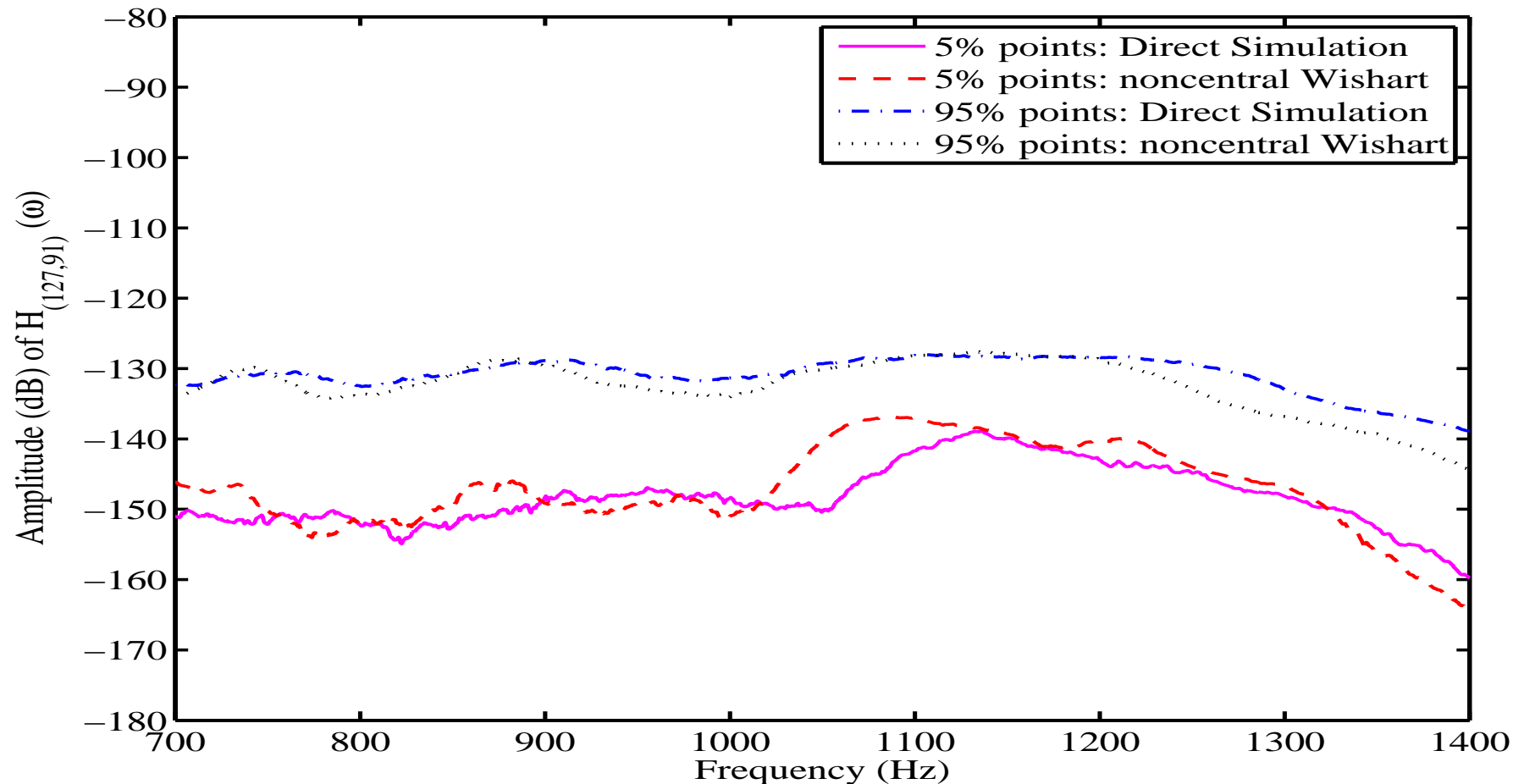
Comparison of cross-FRF: Low Freq



Comparison of the 5% and 95% probability points of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

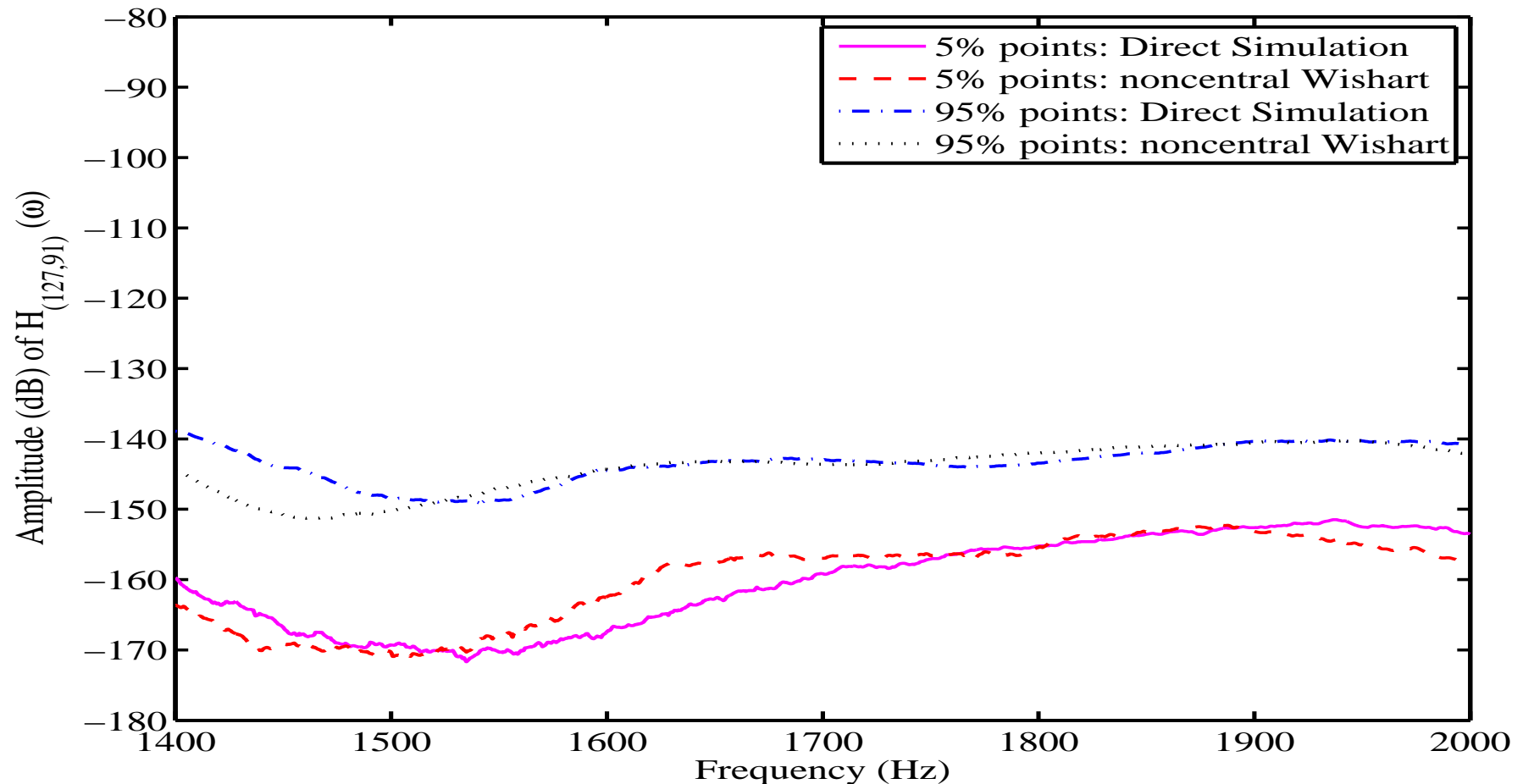
Comparison of cross-FRF: Mid Freq



Comparison of the 5% and 95% probability points of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

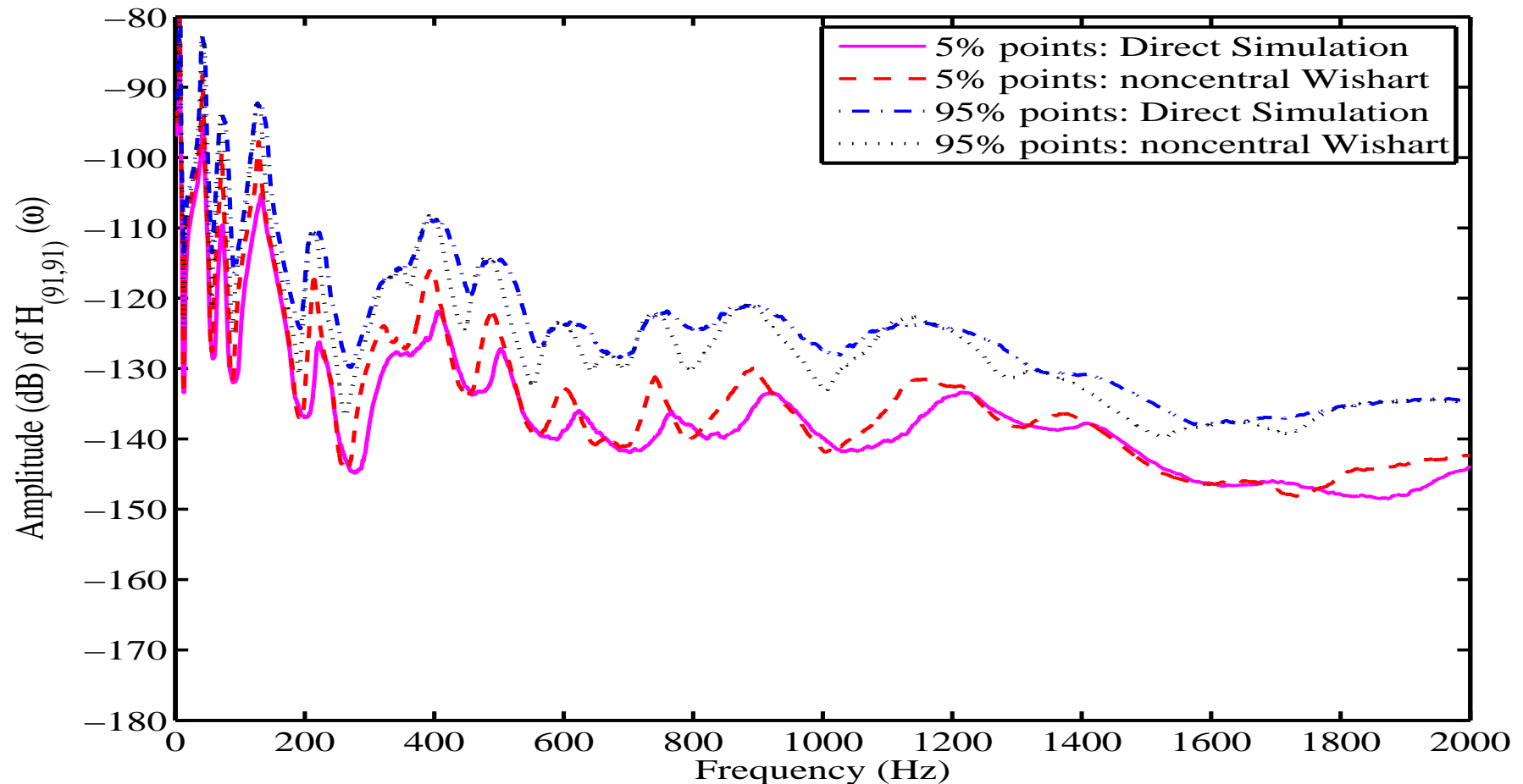
Comparison of cross-FRF: High Freq



Comparison of the 5% and 95% probability points of the amplitude of the cross-FRF, $n = 168$,

$$\delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

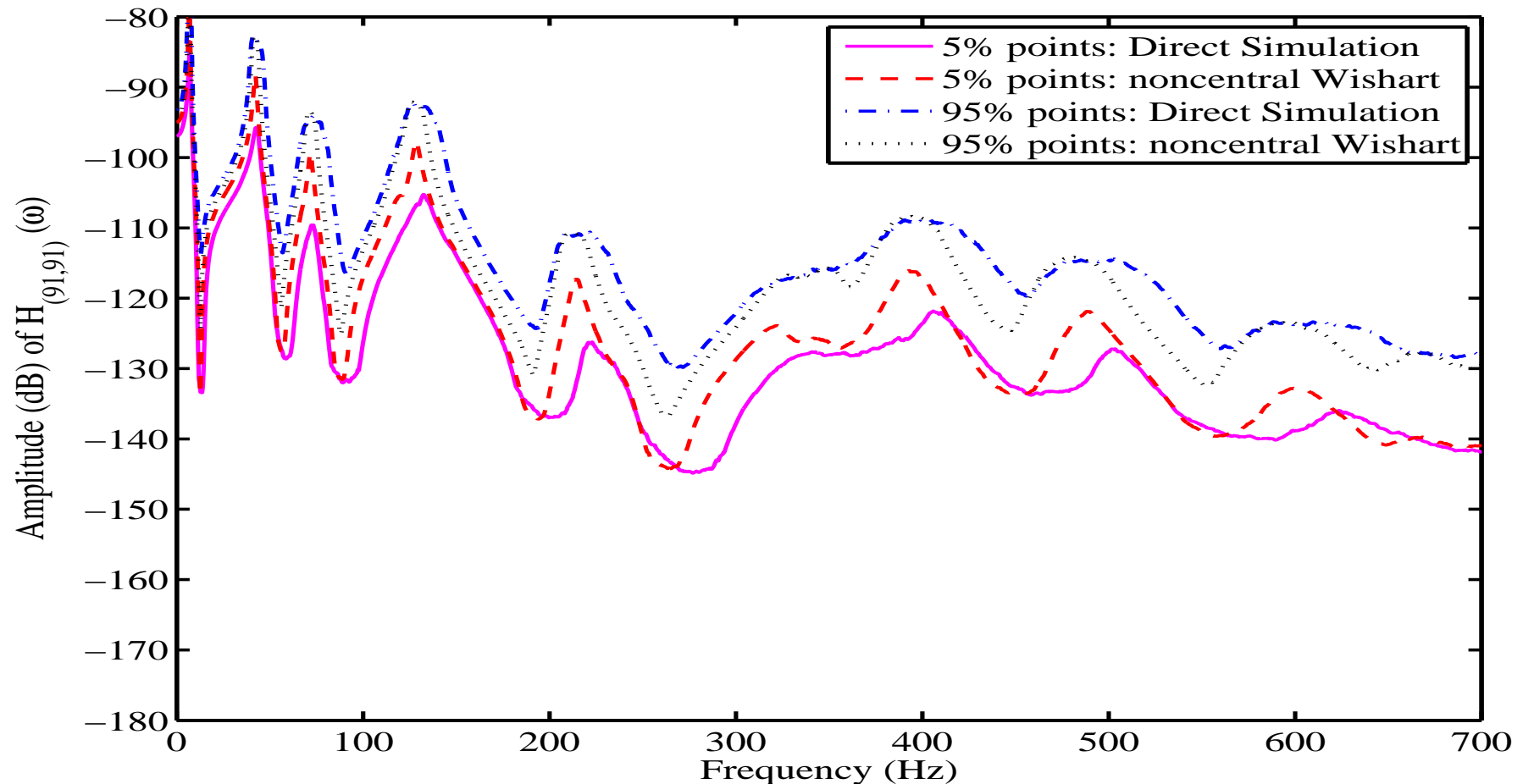
Comparison of driving-point-FRF



Comparison of the 5% and 95% probability points of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

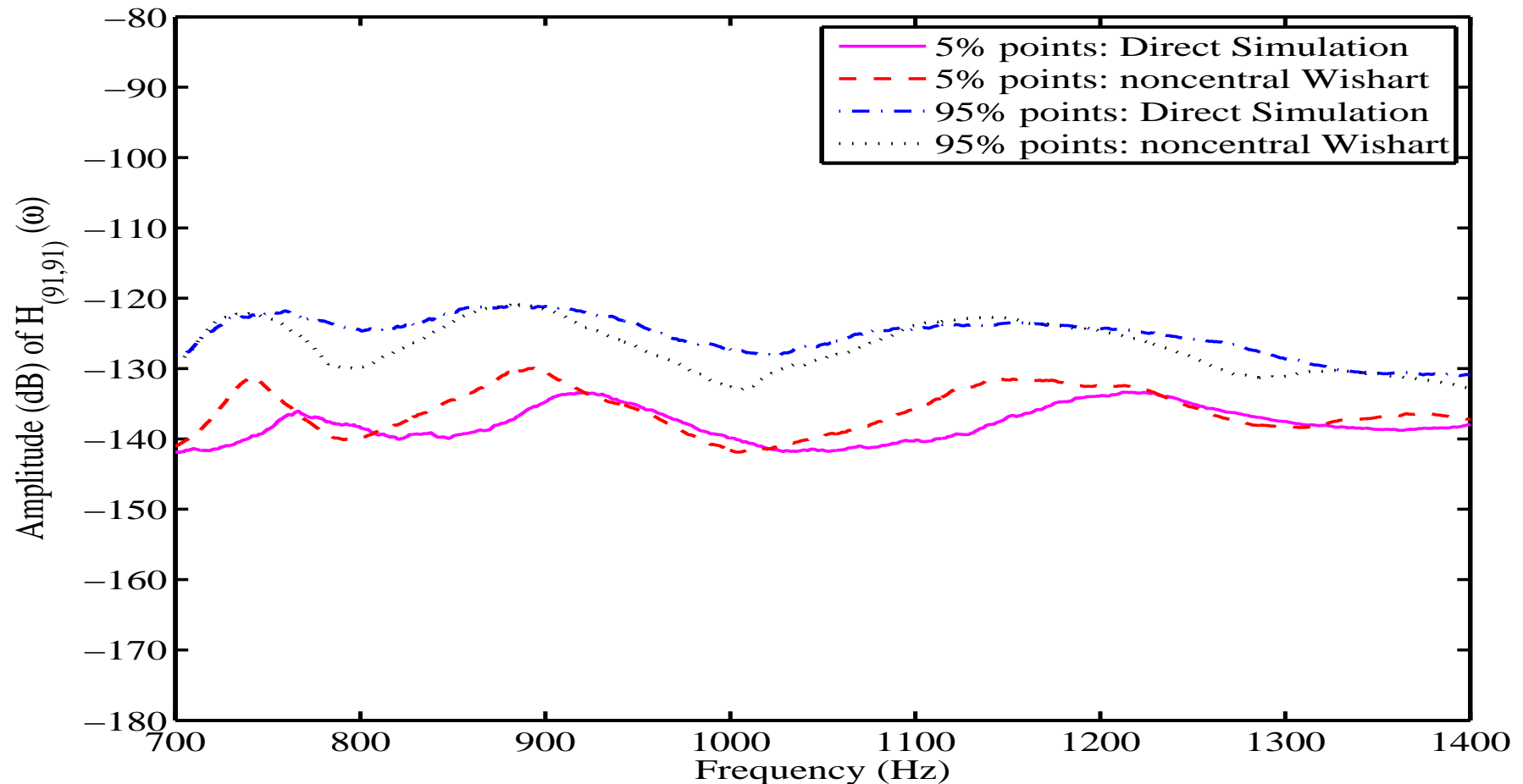
Comparison of driving-point-FRF: Low Freq



Comparison of the 5% and 95% probability points of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

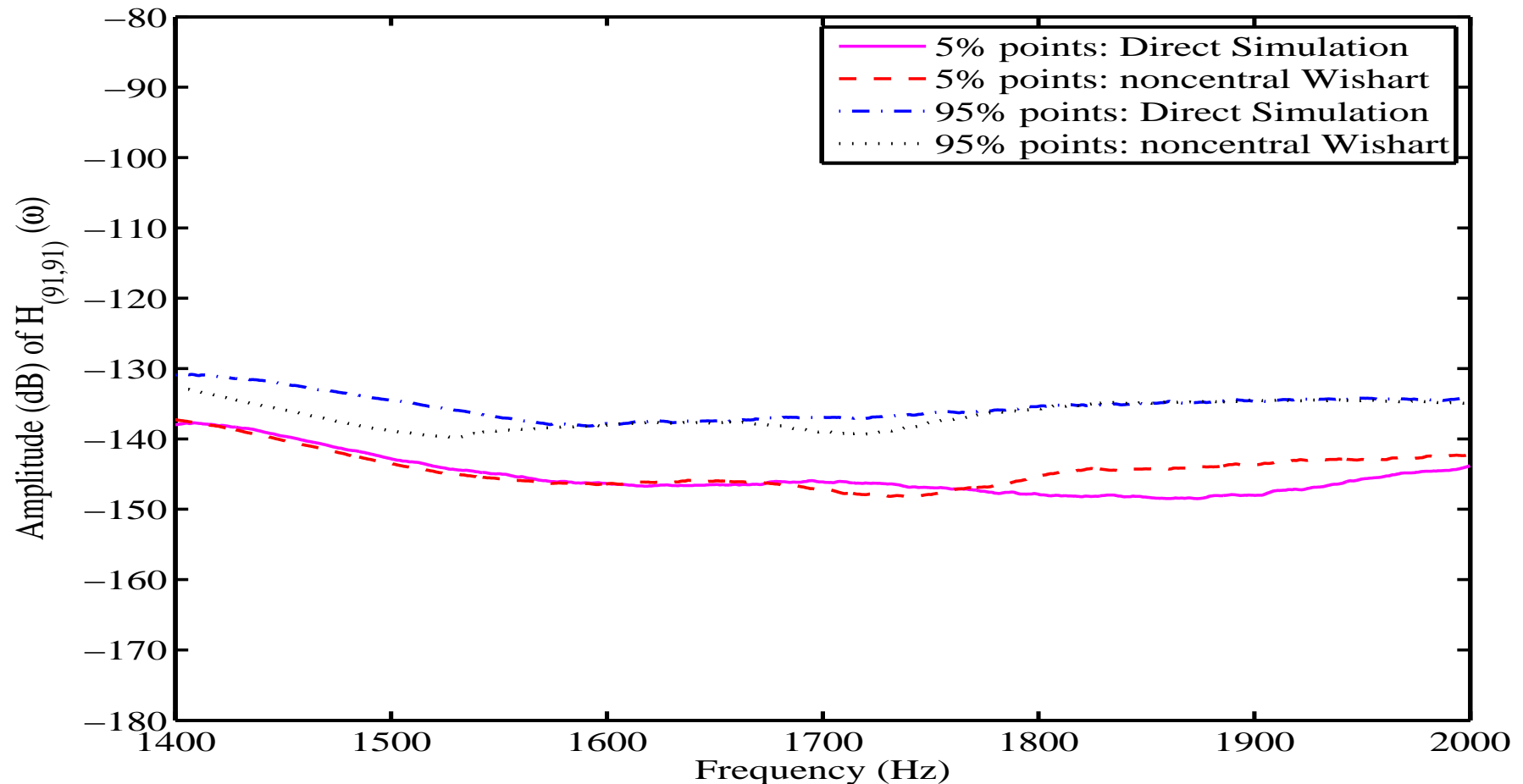
Comparison of driving-point-FRF: Mid Freq



Comparison of the 5% and 95% probability points of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

Comparison of driving-point-FRF: High Freq



Comparison of the 5% and 95% probability points of the amplitude of the driving-point-FRF,

$$n = 168, \delta_M = 0.1166 \text{ and } \delta_K = 0.2711.$$

Conclusions

- When uncertainties in the system parameters (parametric uncertainty) and modelling (nonparametric) are considered, the discretized equation of motion of linear dynamical systems is characterized by random mass, stiffness and damping matrices.
- A new **unified parametric-nonparametric UQ method** for linear dynamical systems has been proposed.
- The matrix variate probability density function of the random system matrices can be represented by noncentral Wishart distribution. Existing nonparametric distribution is a special case of the proposed distribution.

Summary of random matrix models

Random Matrix Model	Comments
<p>Central Wishart/gamma random matrix $W_n(p, \Sigma)$, with $\Sigma = \overline{\mathbf{G}}/p$ (Soize 2001)</p> $p = \frac{1}{\delta_G^2} \left\{ 1 + \frac{\{\text{Trace}(\overline{\mathbf{G}})\}^2}{\text{Trace}(\overline{\mathbf{G}}^2)} \right\} \text{ and}$ $\delta_G^2 = \frac{\mathbb{E}[\ \mathbf{G} - \mathbb{E}[\mathbf{G}]\ _F^2]}{\ \mathbb{E}[\mathbf{G}]\ _F^2} = \frac{\text{Trace}(\mathbf{C}_G)}{\text{Trace}(\overline{\mathbf{G}}^2)}$	<p>(a) The trace of the covariance matrix of the elements of a system matrix is required.</p> <p>(b) The mean of the inverse and the inverse of the mean of the system matrices can be significantly different from each other for the choice of the distribution parameters.</p>
<p>Central Wishart/gamma random matrix $W_n(p, \Sigma)$, with $\Sigma = \overline{\mathbf{G}}/\sqrt{p(p-n-1)}$ and the rest is as defined above (Adhikari 2006).</p>	<p>Parameters are obtained using a least-square error minimization approach. The mean of the matrix and its inverse produce minimum deviations from their respective deterministic values.</p>
<p>Noncentral Wishart random matrix $W_n(p, \Sigma, \Theta)$, with</p> $\Sigma = (\overline{\mathbf{G}} - \Omega)/p, \quad \Theta = \Sigma^{-1}\Omega, \quad p = \frac{\text{Trace}(\overline{\mathbf{G}}^2 - \Omega^2) + \{\text{Trace}(\overline{\mathbf{G}})\}^2 - \{\text{Trace}(\Omega)\}^2}{\delta_G^2 \text{Trace}(\overline{\mathbf{G}}^2)},$ <p>$\Omega \otimes \Omega = \overline{\mathbf{G}} \otimes \overline{\mathbf{G}} - p\mathbf{C}_G/2$ and δ_G is as defined above.</p>	<p>(a) Requires the same information as the previous two distributions</p> <p>(b) If $\Omega = \mathbf{O}_{n,n}$ then this distribution reduces to the central distribution proposed before. The matrix $\Omega \in \mathbb{R}_n^+$ captures the parametric uncertainty through a least-square error minimization involving the covariance matrix \mathbf{C}_G.</p>



FAQs - 1

- *How parametric uncertainties are taken into account?*
- *Since it is a least-square approach, how about the error involved?*
- *Because the covariance matrix is least-square approximated, why not use SFEM as it does not introduce this approximation?*
- *How nonparametric uncertainties are taken into account?*
- *Are you really accounting nonparametric uncertainties?
How do you know ‘unknown unknowns’?*

FAQs - 2

- *I know my uncertainties are localized (e.g., in the joints). Your method introduces uncertainty everywhere in the model. Do you have any recommendations?*
- *How can I use your method if I have no clue about uncertainties in my model?*
- *How much additional computational expense is needed?*
- *How it can be implemented with a commercial FE software?*

FAQs - 3

- *I have never heard of random matrix theory. Is it difficult? Where do I start?*
- *How can I get more information about the unified approach?*