

# Your Title

Your Name

August 3, 2025

Showcase  
Figure

# Outline

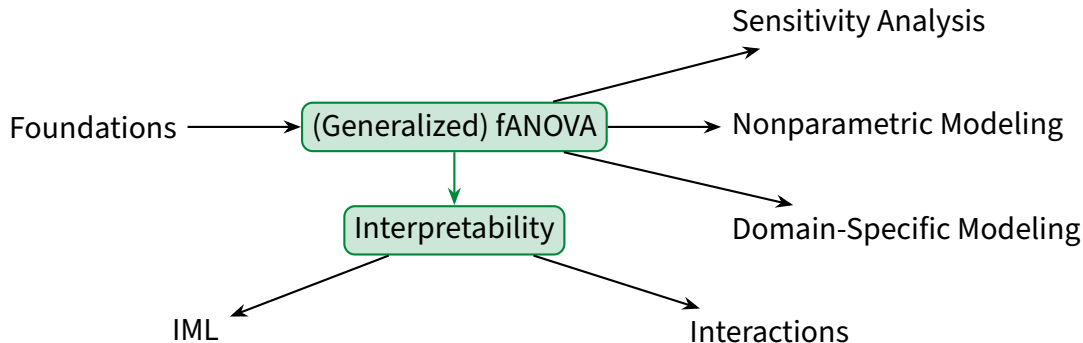
1 Research Context

2 Classical fANOVA

3 Generalized fANOVA

4 Conclusion

5 Extra Slides



References: [1, 2, 5, 7, 6, 4, 3]

# Outline

1 Research Context

2 Classical fANOVA

3 Generalized fANOVA

4 Conclusion

5 Extra Slides

- measure space, probability measure, random vector, subvector, complementary vector, pdf
- measure space of square integrable functions
- inner product
- norm

## General Form

$$\begin{aligned} y(\mathbf{X}) &= \sum_{u \subseteq \{1, \dots, N\}} y_u(\mathbf{X}_u) \\ &= y_{\emptyset} + (y_{\{1\}}(\mathbf{X}_1) + \dots + y_{\{N\}}(\mathbf{X}_N)) \\ &\quad + (y_{\{1,2\}}(\mathbf{X}_1, \mathbf{X}_2) + y_{\{1,3\}}(\mathbf{X}_1, \mathbf{X}_3) + \dots) \\ &\quad + (y_{\{1,2,3\}}(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3) + \dots) + \dots + y_{\{1, \dots, N\}}(\mathbf{X}_1, \dots, \mathbf{X}_N) \end{aligned}$$

- $y$ : Model
- $y_u$ : Component functions for subset  $u$
- Assumption:  $X_1, \dots, X_N$  are independent

## Strong Annihilating Conditions

$$\int_{\mathbb{R}} y_u(\mathbf{x}_u) f_{\{i\}}(x_i) d\nu(x_i) = 0 \quad \text{for } i \in u \neq \emptyset.$$

- Ensures unique component functions
- Applies under independent (product-type) input distributions

$$\mathbb{E}[y_u(\mathbf{x}_u)] = 0$$

$$\mathbb{E}[y_u(\mathbf{x}_u) y_v(\mathbf{x}_v)] = 0 \quad (u \neq v)$$

$$y_{\emptyset} = \int_{\mathbb{R}^N} y(\mathbf{x}) \prod_{i=1}^N f_{\{i\}}(x_i) d\nu(x_i) = \mathbb{E}[y(\mathbf{x})].$$

$$\begin{aligned} y_u(\mathbf{x}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\mathbf{x}_u, \mathbf{x}_{-u}) \prod_{i=1, i \notin u}^N f_{\{i\}}(x_i) d\nu(x_i) - \sum_{v \subsetneq u} y_v(\mathbf{x}_v) \\ &= \mathbb{E}[y(\mathbf{x}_u, \mathbf{x}_{-u}) \mid \mathbf{x}_u = \mathbf{x}_u] - \sum_{v \subsetneq u} y_v(\mathbf{x}_v) \end{aligned}$$

- $f_{-u}$  : marginal density of variables not in  $u$
- Components solved sequentially by increasing order



- $N = 3, u = \emptyset$

$$y_{\emptyset} = \int_{\mathbb{R}^3} y(\mathbf{x}) \prod_{i=1}^3 f_{\{i\}}(x_i) d\nu(x_i) = \mathbb{E}[y(\mathbf{x})].$$

- $u = \{1\} \rightarrow v = \emptyset$

$$y_{\{1\}}(X_1) = \int_{\mathbb{R}^2} y(X_1, x_2, x_3) \prod_{i=2}^3 f_{\{i\}}(x_i) d\nu(x_i) - y_{\emptyset} = \mathbb{E}[y(X_1, X_2, X_3) | X_1 = x_1] - y_{\emptyset}.$$

- $u = \{1, 2\} \rightarrow v = \{1\}, \{2\}, \emptyset$

$$\begin{aligned} y_{\{1,2\}}(X_1, X_2) &= \int_{\mathbb{R}} y(X_1, X_2, x_3) f_{\{3\}}(x_3) d\nu(x_3) - y_{\{1\}}(X_1) - y_{\{2\}}(X_2) - y_{\emptyset} \\ &= \mathbb{E}[y(X_1, X_2, X_3) | X_1 = x_1, X_2 = x_2] - y_{\{1\}}(X_1) - y_{\{2\}}(X_2) - y_{\emptyset}. \end{aligned}$$

$$y(x_1, x_2) = 2x_1 + x_2^2 + x_1x_2$$

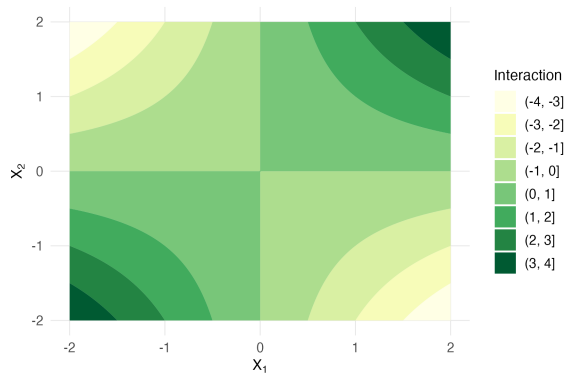
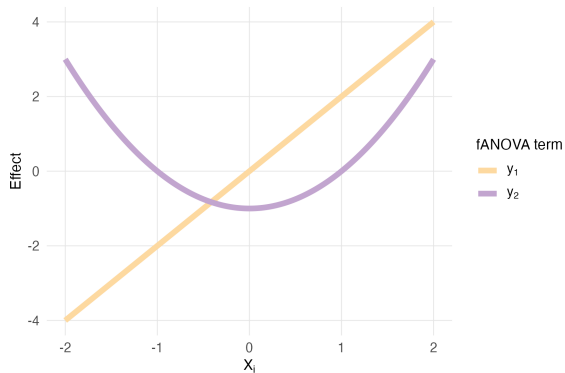
$$(X_1, X_2)^T \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right),$$

$$X_1 \mid X_2 = x_2 \sim \mathcal{N}(0, 1), \quad X_2 \mid X_1 = x_1 \sim \mathcal{N}(0, 1).$$

Components:

$$y_{\emptyset} = 1, \quad y_{\{1\}}(x_1) = 2x_1, \quad y_{\{2\}}(x_2) = x_2^2 - 1, \quad y_{\{1,2\}}(x_1, x_2) = x_1x_2.$$

# Example: 2D Function



$$\Pi_{\mathcal{G}} y = \arg \min_{g \in \mathcal{G}} \|y - g\|^2 = \arg \min_{g \in \mathcal{G}} \mathbb{E}[(y(\mathbf{X}) - g(\mathbf{X}))^2].$$

- $\mathcal{G}$  : linear subspace of  $\mathcal{L}^2$  we project onto
- $g$  all functions in the subspace

$$\begin{aligned}y_{\emptyset} &= \mathbb{E}[y(\mathbf{X})] \\&= \arg \min_{a \in \mathbb{R}} \mathbb{E}[(y(\mathbf{X}) - a)^2] \\&= \arg \min_{g_0 \in \mathcal{G}_0} \|y - g_0\|^2 = \Pi_{\mathcal{G}_0} y,\end{aligned}$$

$$\begin{aligned}y_u(.) &= \mathbb{E}[y(\mathbf{X}) \mid X_u = .] - \sum_{v \subsetneq u} y_v(.) \\&= \arg \min_{g_u \in \mathcal{G}_u} \mathbb{E}[(y(\mathbf{X}) - g_u(.))^2] - \sum_{v \subsetneq u} y_v(.) \\&= (\Pi_{\mathcal{G}_u} y)(.) - \sum_{v \subsetneq u} y_v(.)\end{aligned}$$

## Hoeffding Decomposition

$$y(\mathbf{x}) = \sum_{A \subseteq D} y_A(\mathbf{x}_A), \quad D := \{1, \dots, N\}, \quad (1)$$

where, for each  $A \subseteq D$ , the component function  $y_A$  is defined by:

$$y_A(\mathbf{x}_A) = \sum_{B \subseteq A} (-1)^{|A|-|B|} \mathbb{E}[y(\mathbf{x}) \mid \mathbf{x}_B], \quad (2)$$

where  $y_u$  are orthogonal components.

- Classical fANOVA and Hoeffding decomposition yield same components under zero-centered inputs
- Both assume independence of input variables

$$y(x_1, x_2) = 2x_1 + x_2^2 + x_1x_2$$

$$y_{\emptyset} = \mathbb{E}[y(X_1, X_2)] = 2 \mathbb{E}[X_1] + \mathbb{E}[X_2^2] + \mathbb{E}[X_1X_2] = 1,$$

$$\begin{aligned} y_{\{1\}}(x_1) &= \sum_{B \subseteq \{1\}} (-1)^{1-|B|} \mathbb{E}[y(\mathbf{x}) | X_B] \\ &= -\mathbb{E}[y] + \mathbb{E}[y | X_1 = x_1] \\ &= -1 + (2x_1 + \mathbb{E}[X_2^2] + x_1\mathbb{E}[X_2]) \\ &= 2x_1, \end{aligned}$$

$$\begin{aligned} y_{\{2\}}(x_2) &= \sum_{B \subseteq \{2\}} (-1)^{1-|B|} \mathbb{E}[y(\mathbf{x}) | X_B] \\ &= -\mathbb{E}[y] + \mathbb{E}[y | X_2 = x_2] \end{aligned}$$

$$\begin{aligned}
y_{\{1,2\}}(x_1, x_2) &= \sum_{B \subseteq \{1,2\}} (-1)^{2-|B|} \mathbb{E}[y(\mathbf{x}) | X_B] \\
&= (+1) \mathbb{E}[y] - \mathbb{E}[y | X_1 = x_1] - \mathbb{E}[y | X_2 = x_2] + y(x_1, x_2) \\
&= 1 - (2x_1 + 1) - (x_2^2) + (2x_1 + x_2^2 + x_1x_2) \\
&= x_1x_2.
\end{aligned}$$

$$y(x_1, x_2) = y_{\emptyset} + y_{\{1\}}(x_1) + y_{\{2\}}(x_2) + y_{\{1,2\}}(x_1, x_2) = 1 + 2x_1 + (x_2^2 - 1) + x_1x_2$$



# Outline

1 Research Context

2 Classical fANOVA

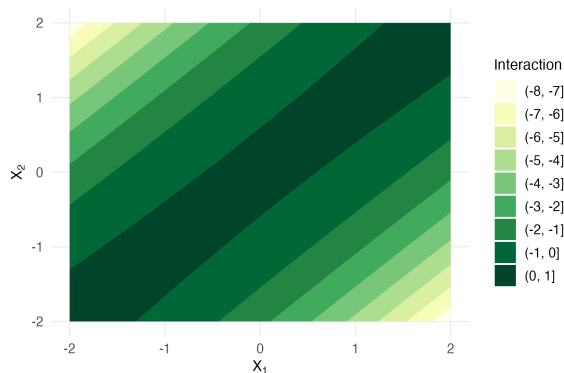
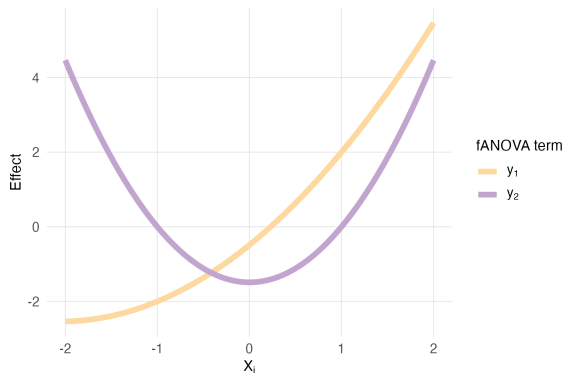
**3 Generalized fANOVA**

4 Conclusion

5 Extra Slides

# Example with Dependent Inputs

$$y(x_1, x_2) = 2x_1 + x_2^2 + x_1x_2, \quad \rho = 0.8$$



## Weak Annihilating Conditions

$$\int_{\mathbb{R}} y_{u,G}(\mathbf{x}_u) f_{\mathbf{x}_u}(\mathbf{x}_u) d\nu(x_i) = 0 \quad \text{for } i \in u \neq \emptyset.$$

- Allows dependent input distributions
- Leads to hierarchical orthogonality

$$\mathbb{E}[y_{u,G}(\mathbf{x}_u)] := \int_{\mathbb{R}^N} y_{u,G}(\mathbf{x}_u) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x}) = 0.$$

$$\mathbb{E}[y_{u,G}(\mathbf{x}_u) y_{v,G}(\mathbf{x}_v)] := \int_{\mathbb{R}^N} y_{u,G}(\mathbf{x}_u) y_{v,G}(\mathbf{x}_v) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x}) = 0.$$

$$y_{\emptyset, G} = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x})$$

$$\begin{aligned} y_{u, G}(\mathbf{x}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\mathbf{x}_u, \mathbf{x}_{-u}) f_{-u}(\mathbf{x}_{-u}) d\nu(\mathbf{x}_{-u}) - \sum_{v \subsetneq u} y_{v, G}(\mathbf{x}_v) \\ &\quad - \sum_{\substack{\emptyset \neq v \subseteq \{1, \dots, N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \int_{\mathbb{R}^{|v \cap u|}} y_{v, G}(\mathbf{x}_{v \cap u}, \mathbf{x}_{v \cap u - u}) f_{v \cap u}(\mathbf{x}_{v \cap u - u}) d\nu(\mathbf{x}_{v \cap u - u}) \end{aligned}$$

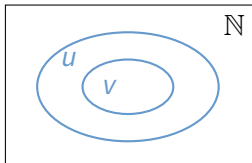
$$y_{\emptyset, G} = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x})$$

$$y_{\emptyset, G} = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x})$$

$$\begin{aligned} y_{u, G}(\mathbf{x}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\mathbf{x}_u, \mathbf{x}_{-u}) f_{-u}(\mathbf{x}_{-u}) d\nu(\mathbf{x}_{-u}) - \sum_{v \subsetneq u} y_{v, G}(\mathbf{x}_v) \\ &\quad - \sum_{\substack{\emptyset \neq v \subseteq \{1, \dots, N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \int_{\mathbb{R}^{|v \cap u|}} y_{v, G}(\mathbf{x}_{v \cap u}, \mathbf{x}_{v \cap u - u}) f_{v \cap u}(\mathbf{x}_{v \cap u - u}) d\nu(\mathbf{x}_{v \cap u - u}) \end{aligned}$$

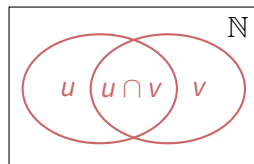
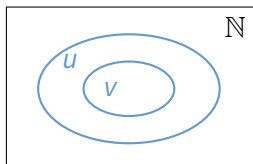
$$y_{\emptyset,G} = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x})$$

$$\begin{aligned} y_{u,G}(\mathbf{x}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\mathbf{x}_u, \mathbf{x}_{-u}) f_{-u}(\mathbf{x}_{-u}) d\nu(\mathbf{x}_{-u}) - \sum_{\substack{v \subsetneq u}} y_{v,G}(\mathbf{x}_v) \\ &\quad - \sum_{\substack{\emptyset \neq v \subseteq \{1, \dots, N\} \\ v \cap u \neq \emptyset, v \not\subseteq u}} \int_{\mathbb{R}^{|v \cap u|}} y_{v,G}(\mathbf{x}_{v \cap u}, \mathbf{x}_{v \cap u}^c) f_{v \cap u}(\mathbf{x}_{v \cap u}^c) d\nu(\mathbf{x}_{v \cap u}^c) \end{aligned}$$



$$y_{\emptyset,G} = \int_{\mathbb{R}^N} y(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x})$$

$$\begin{aligned} y_{u,G}(\mathbf{x}_u) &= \int_{\mathbb{R}^{N-|u|}} y(\mathbf{x}_u, \mathbf{x}_{-u}) f_{-u}(\mathbf{x}_{-u}) d\nu(\mathbf{x}_{-u}) - \sum_{v \subsetneq u} y_{v,G}(\mathbf{x}_v) \\ &\quad - \sum_{\substack{\emptyset \neq v \subseteq \{1, \dots, N\} \\ v \cap u \neq \emptyset, v \not\subset u}} \int_{\mathbb{R}^{|v \cap u|}} y_{v,G}(\mathbf{x}_{v \cap u}, \mathbf{x}_{v \cap u - u}) f_{v \cap u}(\mathbf{x}_{v \cap u - u}) d\nu(\mathbf{x}_{v \cap u - u}) \end{aligned}$$





- $N = 3, u = \emptyset$

$$y_{\emptyset, G} = \mathbb{E}[y(\mathbf{X})]$$

- $u = \{1\} \rightarrow v \subsetneq u = \emptyset$  and  $(\emptyset \neq v \subseteq \{1, \dots, N\}, v \cap u \neq \emptyset, v \not\subseteq u) = \{1, 2\}, \{1, 2, 3\}$

$$\begin{aligned} y_{\{1\}, G}(\mathbf{X}_u) &= \int_{\mathbb{R}^2} y(x_1, x_2, x_3) f_{\{2,3\}}(x_2, x_3) d\nu(x_2, x_3) - y_{\emptyset, G} \\ &\quad - \int_{\mathbb{R}} y_{\{1,2\}, G}(x_1, x_2) f_{\{2\}}(x_2) d\nu(x_2) - \int_{\mathbb{R}} y_{\{1,3\}, G}(x_1, x_3) f_{\{3\}}(x_3) d\nu(x_3) \\ &\quad - \int_{\mathbb{R}^2} y_{\{1,2,3\}, G}(x_1, x_2, x_3) f_{\{2,3\}}(x_2, x_3) d\nu(x_2, x_3) \end{aligned}$$

$$\begin{aligned}y(x_1, x_2) &= a_0 + a_1 x_1 + a_2 x_2 + a_{11} x_1^2 + a_{22} x_2^2 + a_{12} x_1 x_2 \\&= c_0 + c_{1,1} \psi_{1,1}(x_1) + c_{2,1} \psi_{2,1}(x_2) \\&\quad + c_{1,2} \psi_{1,2}(x_1) + c_{2,2} \psi_{2,2}(x_2) + c_{12,11} \psi_{12,11}(x_1, x_2) \\&= \underbrace{c_0}_{y_0} + \underbrace{(c_{1,1} \psi_{1,1}(x_1) + c_{1,2} \psi_{1,2}(x_1))}_{y_1(x_1)} \\&\quad + \underbrace{(c_{2,1} \psi_{2,1}(x_2) + c_{2,2} \psi_{2,2}(x_2))}_{y_2(x_2)} \\&\quad + \underbrace{c_{12,11} \psi_{12,11}(x_1, x_2)}_{y_{12}(x_1, x_2)}.\end{aligned}$$

In [5] Hermite polynomial basis functions are proposed

$$\psi_{\emptyset}(x_1, x_2) = 1,$$

$$\psi_{1,1}(x_1) = x_1,$$

$$\psi_{2,1}(x_2) = x_2,$$

$$\psi_{1,2}(x_1) = x_1^2 - 1,$$

$$\psi_{2,2}(x_2) = x_2^2 - 1,$$

$$\psi_{12,11}(x_1, x_2) = \frac{\rho(x_1^2 + x_2^2)}{1 + \rho^2} - x_1x_2 + \frac{\rho(\rho^2 - 1)}{1 + \rho^2},$$

The corresponding weights can be found via coefficient matching. Start from the interaction term:

$$-c_{12,11} = a_{12} \quad \Rightarrow \quad c_{12,11} = -a_{12}$$

$$c_{1,2} + c_{12,11} \frac{\rho}{1+\rho^2} = a_{11} \quad \Rightarrow \quad c_{1,2} = a_{11} + \frac{\rho}{1+\rho^2} a_{12}$$

$$c_{2,2} + c_{12,11} \frac{\rho}{1+\rho^2} = a_{22} \quad \Rightarrow \quad c_{2,2} = a_{22} + \frac{\rho}{1+\rho^2} a_{12}$$

$$c_{1,1} = a_1$$

$$c_{2,1} = a_2$$

$$c_0 - c_{1,2} - c_{2,2} + c_{12,11} \frac{\rho(\rho^2-1)}{1+\rho^2} = a_0 \quad \Rightarrow \quad c_0 = a_0 + a_{11} + a_{22} + \rho a_{12}$$

- Yields fANOVA components for MVN Inputs
- Works for polynomials of degree up to  $d = 2$

$$\begin{aligned}y_{\emptyset,G} &= a_0 + a_{11} + a_{22} + \rho a_{12}, \\y_{\{1\},G}(x_1) &= a_1 x_1 + \left( a_{11} + \frac{\rho}{1 + \rho^2} a_{12} \right) (x_1^2 - 1), \\y_{\{2\},G}(x_2) &= a_2 x_2 + \left( a_{22} + \frac{\rho}{1 + \rho^2} a_{12} \right) (x_2^2 - 1), \\y_{\{1,2\},G}(x_1, x_2) &= -a_{12} \left( \frac{\rho(x_1^2 + x_2^2)}{1 + \rho^2} - x_1 x_2 + \frac{\rho(\rho^2 - 1)}{1 + \rho^2} \right).\end{aligned}\tag{3}$$

with this Fourier-polynomial expansion, we can build many more examples

In [2] Hooker originally proposed different formulation of generalized fANOVA components:

$$\{y_{u,G}(\mathbf{x}_u) \mid u \subseteq d\} = \arg \min_{\{g_u \in \mathcal{L}^2(\mathbb{R}^{|u|})\}} \int_{\mathbb{R}^N} \left( \sum_{u \subseteq d} g_u(\mathbf{x}_u) - y(\mathbf{x}) \right)^2 f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x}),$$

subject to hierarchical orthogonality conditions:

$$\forall v \subseteq u, \forall g_v : \int_{\mathbb{R}^N} y_u(\mathbf{x}_u) g_v(\mathbf{x}_v) f_{\mathbf{x}}(\mathbf{x}) d\nu(\mathbf{x}) = 0.$$

$$\mu := \mathbb{E}[y(\mathbf{X})] = y_{\emptyset, G}.$$

$$\begin{aligned}\sigma^2 &:= \mathbb{E} \left[ (y(\mathbf{X}) - \mu)^2 \right] \\ &= \mathbb{E} \left[ \left( y_{\emptyset, G} + \sum_u y_{u, G}(\mathbf{x}_u) - y_{\emptyset, G} \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \sum_u y_{u, G}(\mathbf{x}_u) \right)^2 \right] \\ &= \sum_u \mathbb{E} [y_{u, G}^2(\mathbf{x}_u)] + \sum_{u \not\subseteq v, v \not\subseteq u} \mathbb{E} [y_{u, G}(\mathbf{x}_u) y_{v, G}(\mathbf{x}_v)],\end{aligned}$$



# Outline

- 1 Research Context
- 2 Classical fANOVA
- 3 Generalized fANOVA
- 4 Conclusion**
- 5 Extra Slides



- Estimation schemes and software implementation
- Extension of Fourier polynomial expansion to other distributions

# Outline

1 Research Context

2 Classical fANOVA

3 Generalized fANOVA

4 Conclusion

5 Extra Slides

- Zero mean property: factorized density, Fubini's Theorem, strong annihilating conditions
- Mutual orthogonality: factorized density, Fubini's Theorem, strong annihilating conditions

- Zero mean property: separating  $x$  into subvectors, marginal density, Fubini's Theorem, weak annihilating conditions
- Hierarchical orthogonality: set the scene,  $u$  is a proper subset of  $v$   $u \subsetneq v$ , so there is an index in  $u$  which is not in  $v$ ; divide  $x_u$  into subvectors, marginal density, Fubini and weak annihilating conditions
- Weak annihilating becomes strong under independence: assume the weak ones, product density, factor out
- Three integration cases: distinguish between different relationships  $u$  and  $v$ , depending on the relationship the integral w.r.t. to marginal density simplifies
- Generalized fANOVA components by Rahman: first build constant term; for nonconstant terms use integration cases
- Integration constraint Hooker: show that hierarchical orthogonality is fulfilled if the conditions hold, show that it is not fulfilled if they do not hold; but why exactly these conditions a bit unclear
- Take a look at Sobol's proof again

- [https://docs.google.com/spreadsheets/d/1K5ECL6hDPDnHwM\\_k342xa29H-vHWzdk27PTgDHUwfFE/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1K5ECL6hDPDnHwM_k342xa29H-vHWzdk27PTgDHUwfFE/edit?usp=sharing) - Table with fANOVA-related literature



Giles Hooker.

Discovering additive structure in black box functions.

*In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 575–580, Seattle WA USA, August 2004. ACM.



Giles Hooker.

Generalized Functional ANOVA Diagnostics for High-Dimensional Functions of Dependent Variables.

*Journal of Computational and Graphical Statistics*, 16(3):709–732, September 2007.



Gunnar König, Eric Günther, and Ulrike von Luxburg.

Disentangling Interactions and Dependencies in Feature Attribution, October 2024.  
[arXiv:2410.23772](https://arxiv.org/abs/2410.23772).





Benjamin Lengerich, Sarah Tan, Chun-Hao Chang, Giles Hooker, and Rich Caruana.  
Purifying Interaction Effects with the Functional ANOVA: An Efficient Algorithm for  
Recovering Identifiable Additive Models.

*In Proceedings of the Twenty Third International Conference on Artificial Intelligence and  
Statistics*, pages 2402–2412. PMLR, June 2020.

ISSN: 2640-3498.



Sharif Rahman.

A Generalized ANOVA Dimensional Decomposition for Dependent Probability Measures.

*SIAM/ASA Journal on Uncertainty Quantification*, 2(1):670–697, 2014.



I. M. Sobol.

Sensitivity Estimates for Nonlinear Mathematical Models.

*Mathematical Modelling and Computational Experiments*, 1:407–414, 1993.



Charles J. Stone.

The Use of Polynomial Splines and Their Tensor Products in Multivariate Function Estimation.

*The Annals of Statistics*, 22(1):118–171, 1994.

Publisher: Institute of Mathematical Statistics.