

# Содержание

<b>1</b>	<b>Описание стендового оборудования и сбор данных</b>	<b>5</b>
1.1	Состав стенда . . . . .	5
1.2	Параметры измерений и методика сбора данных . . . . .	6
1.3	Методика извлечения признаков во временной области . . . . .	6
1.4	Методика извлечения признаков в частотной области . . . . .	6
<b>2</b>	<b>Обучение модели и анализ важности признаков</b>	<b>8</b>
2.1	Описание алгоритма случайного леса . . . . .	8
2.2	Настройка и обучение модели . . . . .	8
2.3	Оценка важности признаков . . . . .	8
2.4	Результаты и обсуждение . . . . .	9

# Примерное содержание

1. Рассматриваются этапы подготовки данных, выбор и обучение модели, а также оценка её качества.
2. В качестве данных используются временные ряды, полученные с датчиков тока асинхронного электродвигателя.
  - (a) Приводится описание и структурная схема стендового оборудования, предназначенного для сбора данных.
  - (b) Описываются режимы работы электропривода, виды имитируемых неисправностей, а также процесс сбора данных (частота дискретизации, длительность записи и т.д.).
  - (c) Описываются методы предварительной обработки данных, включая фильтрацию шумов.
3. Приводится методика извлечения признаков из временных рядов, включая статистические характеристики, и частотные признаки.
  - (a) Описываются алгоритмы выделения признаков, такие как преобразование Фурье (но с использованием welch), разбиение временных признаков на окна, разбиение частотной области на полосы, приводится спектрограмма.
  - (b) Приводятся некоторые формулы для расчета признаков.
4. Рассматриваются процесс обучения модели случайного леса, включая выбор гиперпараметров и методы кросс-валидации. Рассматривается вычисления важности признаков методами SHAP, MDI и Permutation Importance.
  - (a) Приводится описание алгоритма случайного леса, его преимущества и недостатки.
  - (b) Приводятся результаты обучения модели, включая метрики качества (точность, полнота, F1-мера и т.д.).
  - (c) Приводятся графики важности признаков, производится переобучение модели на отобранных признаках.
  - (d) Обсуждаются результаты, включая интерпретацию важности признаков, возможный вклад и физический смысл наиболее значимых признаков. Сравниваются результаты методов оценки важности признаков.
5. Приводится заключение, включая основные выводы и направления для будущих исследований.

# Введение

Асинхронные двигатели находят широкое применение в различных отраслях промышленности, таких как металлургия, нефтехимия, водоснабжение, транспорт и производство. Например, они используются для привода насосов, вентиляторов, конвейеров, компрессоров и других механизмов, требующих надежного и эффективного электропривода. Несмотря на высокую надежность, асинхронные двигатели подвержены различным неисправностям. Наиболее распространенной причиной отказов являются повреждения подшипников, на долю которых приходится до половины всех случаев отказа электродвигателей [1]. Кроме того, встречаются электрические неисправности (межвитковые замыкания, пробой изоляции, обрыв фазы и т.д.), а также механические дефекты (дисбаланс, перекос ротора и т.д.) и нарушения в питающей сети. Эти неисправности могут привести к простою производства, снижению качества продукции и значительным финансовым потерям. В связи с этим особое значение приобретает применение современных и эффективных методов мониторинга состояния, позволяющих своевременно выявлять и диагностировать неисправности асинхронных электродвигателей.

В настоящее время реактивное обслуживание постепенно заменяется мониторингом состояния и предсказательным обслуживанием, что позволяет автоматизировать оценку состояния электродвигателя и снизить зависимость от трудоёмких ручных проверок. Совершенствование методов мониторинга и диагностики способствует раннему выявлению отклонений и предотвращению простоев оборудования, что, в свою очередь, уменьшает затраты на обслуживание и повышает надёжность системы. По этой причине мониторинг состояния электроприводов получил значительное внимание в научной литературе [2–5].

Существуют два основных подхода к диагностике состояния электроприводов: модельно-ориентированный и основанный на данных. Модельно-ориентированный подход использует математические модели, описывающие физические процессы и динамику системы. Такие методы могут обеспечивать высокую точность, однако требуют глубоких знаний о конструкции, параметрах и принципах работы оборудования, что ограничивает их применение в реальных промышленных условиях. С другой стороны, методы, основанные на данных, включают современные технологии искусственного интеллекта, такие как искусственные нейронные сети (ИНС), классические алгоритмы машинного обучения и методы нечеткой логики. Эти подходы приобрели широкую популярность в последние годы и особенно эффективны, когда имеется доступ к данным с датчиков, но математическая модель системы неизвестна или слишком сложна для построения. Алгоритмы машинного обучения позволяют решать задачи классификации и прогнозирования состояния оборудования, обеспечивая высокую гибкость и адаптивность к различным типам неисправностей.

В задачах мониторинга и диагностики состояния электроприводов наибольшее распространение получили анализ вибрационных сигналов и фазных токов. Вибрационные данные традиционно применяются для выявления механических неисправностей, в то время как анализ фазных токов эффективен как для обнаружения электрических дефектов, так и для косвенной оценки механических, поскольку они отражаются на электромагнитных процессах. Для обработки таких сигналов широко используются методы анализа во временной, частотной и временно-частотной областях (например, преобразование Фурье, вейвлет-преобразование и преобразование Хильберта–Хуанга). Подходы MCSA (Motor Current Signature Analysis) и вибрационного анализа (Vibration Signature Analysis, VSA) основаны на выделении информативных спектральных признаков, характеризующих состояние привода. Помимо вибрации и токов, дополнительными источниками информации могут служить температура, скорость вращения, акустические сигналы и параметры питающей сети.

Для снижения влияния шума и выделения информативных признаков, отражающих рабочие и неисправные состояния электропривода, выполняется предварительная обработка сигналов. На этом этапе могут применяться методы заполнения пропусков, фильтрации, нормализации и другие процедуры улучшения качества данных. Полученные признаки формируют основу для построения модели классификации и последующего анализа состояния оборудования.

Для повышения эффективности и интерпретируемости модели применяется отбор признаков (feature selection). Он позволяет исключить избыточные и коррелированные параметры, сократить размерность пространства признаков и тем самым повысить устойчивость моделей к переобучению, улучшить её обобщающую способность и уменьшить вычислительные затраты. В качестве фильтрационных подходов могут использоваться порог дисперсии и анализ кросс-корреляции между признаками; для снижения размерности — метод главных компонент (Principal Component Analysis, PCA). Оценка значимости признаков осуществляется с помощью современных методов интерпретации, таких как MDI (Mean Decrease in Impurity), Permutation Importance и SHAP values.

После формирования оптимального набора признаков осуществляется построение модели классификации, предназначенной для распознавания состояния электропривода. В задачах технической диагностики применяются как классические алгоритмы машинного обучения, так и современные методы глубокого обучения. К наиболее распространённым относятся Random Forest, Support Vector Machine (SVM), k-Nearest Neighbors (kNN), Gradient Boosting и различные модификации нейронных сетей. Выбор конкретной модели определяется особенностями исходных данных, объёмом обучающей выборки и требованиями к интерпретируемости и вычислительным ресурсам. [ОБЗОР СТАТЕЙ]

Для оценки качества классификаторов обычно применяются стандартные метрики — точность (accuracy), полнота (recall), точность положительных предсказаний (precision) и интегральный показатель F1-score.

Цель данной работы — оценить возможности диагностики механических неисправностей асинхронного электродвигателя путем анализа фазных токов. Экспериментальные данные, представляющие собой временные ряды фазных токов в разных режимах работы и при различных неисправностях, были получены на экспериментальном стенде. В работе описан процесс их обработки, а также формирования информативных признаков в двух областях: временной и частотной. На основе извлечённых признаков была обучена модель классического машинного обучения Random Forest (RF), и проведена оценка точности её классификации по основным метрикам. Центральное место в исследовании занял сравнительный анализ важности признаков, проведенный тремя методами (SHAP, MDI, Permutation Importance). Это позволило не только подтвердить эффективность подхода, но и выявить набор наиболее значимых признаков, напрямую характеризующих механическое состояние электропривода.

# 1. Описание стендового оборудования и сбор данных

## 1.1. Состав стенда

Структурная схема стенда для исследования асинхронного электропривода приведена на рисунке 1.1.

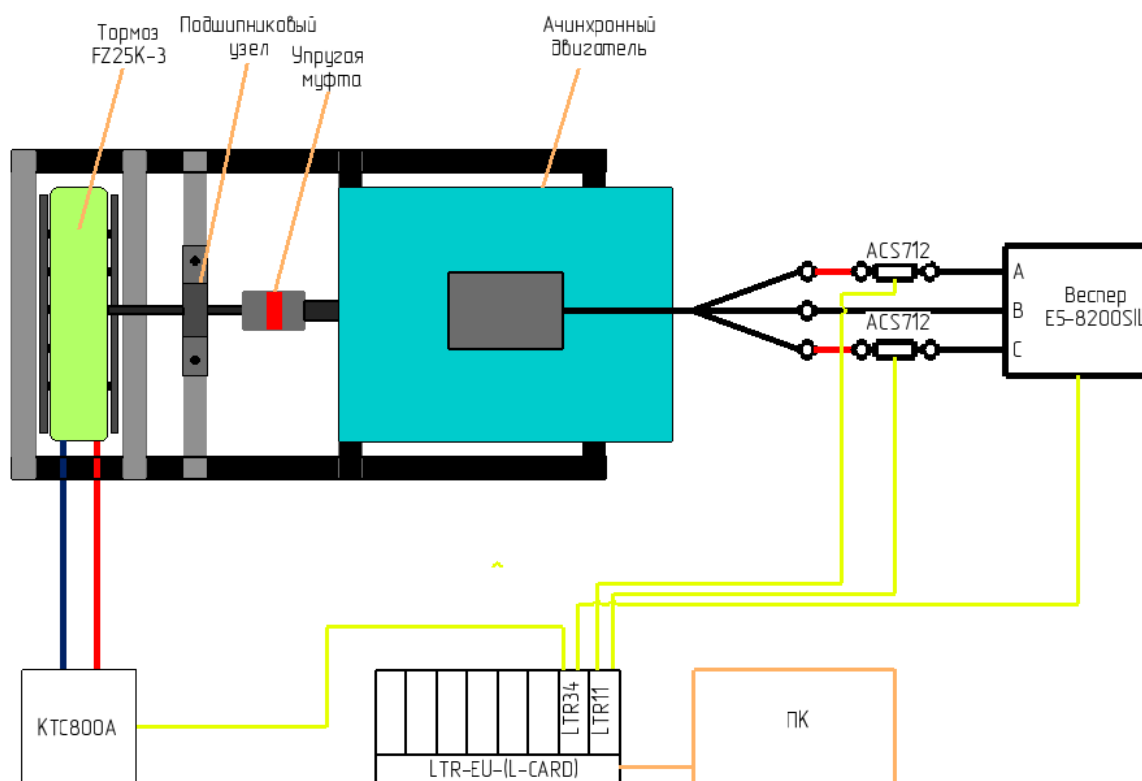


Рис. 1.1: Структурная схема стендового оборудования для исследования асинхронного электропривода

В составе стенда входит следующее основное оборудование:

- Асинхронный двигатель ЕЛПРОМ ТРОЯН 0,75кВт
- Тормоз FZ25K-3
- Подшипниковый узел
- Упругая муфта
- Частотный преобразователь Веспер ES-8200SIL
- Датчики тока ACS712

Объектом исследования является асинхронный двигатель. Он подключен к частотному преобразователю, который обеспечивает питание двигателя переменным током с регулируемой частотой. Тормоз используется для создания нагрузки на вал двигателя, что позволяет моделировать различные режимы работы. Упругая муфта и подшипниковый узел соединяют вал двигателя с тормозом, позволяя передавать крутящий момент при различных

условиях нагрузки. Датчики тока ACS712 установлены на фазах статора двигателя для измерения токов в реальном времени.

## 1.2. Параметры измерений и методика сбора данных

Для проведения исследований был проведен сбор экспериментальных данных сигналов фазного тока асинхронного двигателя при различных технических состояниях. Регистрация двух фаз тока статора (каналы А и В) осуществлялась с частотой дискретизации 10 кГц, при этом длительность каждой записи составляла 10 с, что соответствует 100 000 отсчетов на каждый зарегистрированный сигнал. Испытания проводились при электрической частоте питания 20 Гц в двух режимах нагрузки: холостой ход и нагрузка, соответствующая 50% от номинальной мощности двигателя.

Было исследовано четыре диагностируемых состояния электродвигателя. **Нормальная работа** соответствовала исправному двигателю, отключенному от муфты стенда. Состояние **дисбаланса ротора** достигалось подключением двигателя к муфте, масса которой создавала неуравновешенность. **Перекося ротора (несоосность)** имитировался установкой дополнительной шайбы толщиной 0,3 мм под опорным подшипниковым узлом. **Износ подшипников** моделировался путем установки подшипника с искусственным дефектом на внешнем кольце в двигатель.

Для обеспечения статистической достоверности каждый эксперимент повторялся многократно (табл. ??). Испытания проводились по фиксированному протоколу: разгон двигателя до установившегося режима в течение 10-15 секунд, запись сигнала длительностью 10 с и сохранение метаданных эксперимента (класс дефекта, режим работы, дата/время).

Для увеличения объема выборки и приведения данных к единому размеру, необходимому для последующего анализа, каждый зарегистрированный сигнал подвергался процедуре сегментации. Исходные временные ряды разбивались на последовательные отрезки с использованием скользящего окна длиной 4096 отсчетов и перекрытием 50% между соседними сегментами. Данная процедура позволила сформировать массив стандартизированных фрагментов сигналов для дальнейшей обработки.

## 1.3. Методика извлечения признаков во временной области

Для каждого сегмента сигнала фазного тока, полученного в процессе сегментации, вычислялся набор статистических признаков, отражающих форму, энергию и вариативность сигнала во временной области. Эти характеристики позволяют количественно описывать основные свойства сигнала без перехода в частотную область и широко применяются в задачах диагностики технического состояния электрических машин. В работе были использованы следующие признаки: среднееквадратичное значение, коэффициент асимметрии, коэффициент эксцесса, размах, коэффициент импульсности и коэффициент формы [1].

## 1.4. Методика извлечения признаков в частотной области

Для анализа частотных характеристик сигналов фазных токов использовались методы спектральной и огибающей обработки. Предварительный анализ амплитудно-частотных спектров показал наличие выраженных гармонических составляющих, обусловленных широтно-импульсной модуляцией (ШИМ) инвертора. На спектрограммах (рис. ??) отчетливо наблюдались две доминирующие гармоники несущей на частотах около **1665 Гц** и **3330 Гц**, не связанных с механическим состоянием двигателя. Их присутствие маскировало диагностически значимые низкочастотные составляющие, что делало прямое применение преобразования Фурье неинформативным.

В связи с этим для выделения информативной низкочастотной компоненты огибающей тока была применена методика демодуляции с использованием преобразования Гильберта. Обработка каждого сегмента сигнала включала следующие этапы:

Обработка каждого сегмента сигнала включала следующие этапы:

1. **Полосовая фильтрация.** Сигнал  $x(t)$  пропусклся через полосовой фильтр Баттерворта четвертого порядка с центральной частотой  $f_c = 1670$  Гц и полосой пропускания  $\pm 50$  Гц:

$$x_f(t) = \text{BPF}\{x(t)\}, \quad 1620 \text{ Гц} \leq f \leq 1720 \text{ Гц}.$$

2. **Преобразование Гильберта.** Для вычисления мгновенной амплитуды (огибающей) формировался аналитический сигнал:

$$z(t) = x_f(t) + j \mathcal{H}\{x_f(t)\},$$

где  $\mathcal{H}\{\cdot\}$  — преобразование Гильберта. Огибающая сигнала определяется как:

$$e(t) = |z(t)| = \sqrt{x_f^2(t) + \mathcal{H}^2\{x_f(t)\}}.$$

3. **Фильтрация низких частот.** Для устранения остаточных высокочастотных колебаний применялся фильтр нижних частот (ФНЧ) с граничной частотой  $f_{LP} = 200$  Гц, обеспечивающий выделение диагностически значимых низкочастотных составляющих огибающей.
4. **Спектральный анализ.** Для оценки спектральной плотности мощности огибающего сигнала использовался метод Вэлча с окном Ханна длиной 512 отсчётов и перекрытием 50% (256 отсчётов). Итоговая оценка спектральной плотности мощности вычислялась по формуле:

$$P_{xx}(f) = \frac{1}{K} \sum_{k=1}^K \frac{|\text{FFT}\{w(n)x_k(n)\}|^2}{U},$$

где  $w(n)$  — окно Ханна,  $U$  — нормировочный коэффициент энергии окна,  $K$  — число сегментов.

5. **Извлечение признаков.** На основе спектра огибающей вычислялся набор диагностических признаков: **(Взять из наиболее значимых!)**

## 2. Обучение модели и анализ важности признаков

### 2.1. Описание алгоритма случайного леса

Для классификации технического состояния асинхронного двигателя использовался алгоритм **случайного леса** (Random Forest), основанный на ансамблевом подходе объединения множества независимых решающих деревьев. Каждое дерево обучается на случайно выбранной подвыборке обучающих данных (bagging) и случайном подмножестве признаков, что снижает вероятность переобучения и повышает обобщающую способность модели.

Модель случайного леса характеризуется рядом преимуществ: устойчивостью к шуму и выбросам, возможностью работы с нелинейными зависимостями, а также встроенными механизмами оценки важности признаков. К недостаткам можно отнести сравнительно большую вычислительную сложность при построении ансамбля и снижение интерпретируемости по сравнению с отдельным решающим деревом.

### 2.2. Настройка и обучение модели

Перед обучением проводилось масштабирование и нормализация признаков, после чего выборка данных разбивалась на обучающую и тестовую подвыборки в соотношении 80/20. Для повышения достоверности оценки качества классификации использовалась **стратифицированная кросс-валидация** с числом блоков  $k = 5$ .

Обучение выполнялось с подбором основных гиперпараметров модели (число деревьев, максимальная глубина, минимальное число выборок в узле и т.д.) на основе средних результатов кросс-валидации по метрике точности (Accuracy).

Метрики качества классификации включали:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad \text{Precision} = \frac{TP}{TP + FP}, \quad \text{Recall} = \frac{TP}{TP + FN},$$

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}},$$

где  $TP, TN, FP, FN$  — количество истинных и ложных классификаций по классам.

### 2.3. Оценка важности признаков

Для интерпретации результатов классификации и анализа вклада отдельных признаков в процесс принятия решений использовались три метода оценки важности:

- **MDI (Mean Decrease Impurity)** — среднее уменьшение неопределённости в узлах деревьев (встроенная метрика модели Random Forest).
- **Permutation Importance** — оценка снижения точности модели при случайном перемешивании значений каждого признака.
- **SHAP (SHapley Additive exPlanations)** — метод на основе теории Шепли, позволяющий количественно определить вклад каждого признака в предсказание модели.

Визуализация важности признаков для различных методов представлена на рисунках ??–??. Наиболее значимые признаки интерпретировались с точки зрения их физического смысла, что позволило установить корреляцию между вычисленными показателями и типами дефектов.



## 2.4. Результаты и обсуждение

Результаты обучения модели показали высокую точность классификации различных технических состояний электродвигателя. Анализ важности признаков позволил выявить наиболее информативные показатели, отражающие физические процессы в системе.

Таблица 2.1: Результаты анализа признаков для здорового двигателя

Ранг	Признак	Среднее	СКО	Мин	Макс
1	<b>СКЗ тока фазы</b> <i>current_phase_a_rms</i>	0,244	0,003	0,239	0,249
2	<b>Размах огибающей тока</b> <i>current_phase_a_env_ptp</i>	0,040	0,010	0,029	0,071
3	<b>Показатель регулярности гармоник огибающей тока фазы</b> <i>current_phase_a_env_harmonic_regularity</i>	0,979	0,017	0,938	1,000
4	<b>Дисбаланс энергии фаз А-Б</b> <i>current_phase_a_current_phase_b_env_energy_imbalance</i>	0,145	0,141	0,003	0,524
5	<b>Коэффициент формы</b> <i>current_phase_a_env_form_factor</i>	1,134	0,021	1,090	1,217

Таблица 2.2: Результаты анализа признаков для неисправного подшипника

Ранг	Признак	Среднее	СКО	Мин	Макс
1	<b>СКЗ тока фазы</b> <i>current_phase_a_rms</i>	0,244	0,003	0,238	0,256
2	<b>Размах огибающей тока</b> <i>current_phase_a_env_ptp</i>	0,043	0,016	0,029	0,110
3	<b>Гармоническая составляющая мощности огибающей</b> <i>current_phase_a_env_harmonic_ratio</i>	0,837	0,168	0,291	1,000
4	<b>Дисбаланс энергии фаз А-Б</b> <i>current_phase_a_current_phase_b_env_energy_imbalance</i>	0,167	0,163	0,000	0,613
5	<b>Энтропия Винера</b> <i>current_phase_a_env_flatness</i>	$2,49 \times 10^{-5}$	$6,74 \times 10^{-6}$	$7,98 \times 10^{-6}$	$4,76 \times 10^{-5}$

## Обсуждение результатов

Анализ статистических характеристик признаков, представленных в таблицах 2.1–2.4, позволяет выделить ключевые тенденции, отражающие различия токовых сигналов при разных состояниях электродвигателя.

**1. СКЗ тока фазы (RMS).** Данный показатель стабильно занимает первое место по значимости для всех состояний. Для случая несоосности наблюдается наибольшее среднее значение (0,275 против 0,244 у здорового двигателя), что связано с повышением механических потерь и увеличением нагрузки на вал. При ”несоосности стэнда”, напротив, среднее значение RMS снижается до 0,224, что может указывать на ???

**2. Размах огибающей тока.** Размах огибающей (*Envelope PTP*) увеличивается как при неисправности подшипника, так и при несоосности, отражая появление модуляций тока на частотах механических дефектов. Для подшипникового дефекта наблюдается наиболее широкий диапазон (0,029–0,110), что свидетельствует о неустойчивости момента и пульсациях при вращении ротора.

Таблица 2.3: Результаты анализа признаков для несоосности

Ранг	Признак	Среднее	СКО	Мин	Макс
1	<b>СКЗ тока фазы</b> <i>current_phase_a_rms</i>	0,275	0,005	0,267	0,287
2	<b>Размах огибающей тока</b> <i>current_phase_a_env_ptp</i>	0,040	0,010	0,030	0,092
3	<b>Показатель регулярности гармоник огибающей тока фазы</b> <i>current_phase_a_env_harmonic_regularity</i>	0,969	0,018	0,938	1,000
4	<b>Спектральная корреляция фаз А-Б</b> <i>current_phase_a_current_phase_b_env_spectral_corr</i>	0,911	0,088	0,559	0,992
5	<b>Доля мощности гармонических искажений (THD) огибающей тока фазы</b> <i>current_phase_a_env_thd_power_frac</i>	0,438	0,101	0,203	0,589

Таблица 2.4: Результаты анализа признаков для несоосности стэнда

Ранг	Признак	Среднее	СКО	Мин	Макс
1	<b>СКЗ тока фазы А</b> <i>current_phase_a_rms</i>	0,224	0,004	0,218	0,238
2	<b>Энтропия Винера</b> <i>current_phase_a_env_flatness</i>	$2,91 \times 10^{-5}$	$1,29 \times 10^{-5}$	$1,59 \times 10^{-5}$	$9,15 \times 10^{-5}$
3	<b>Коэффициент асимметрии</b> <i>current_phase_a_env_skew</i>	0,409	0,663	-0,344	3,372
4	<b>Дисбаланс энергии фаз А-Б</b> <i>current_phase_a_current_phase_b_env_energy_imbalance</i>	0,165	0,145	0,005	0,567
5	<b>Спектральная корреляция фаз А-Б</b> <i>current_phase_a_current_phase_b_env_spectral_corr</i>	0,903	0,077	0,591	0,991

**3. Гармоническая структура огибающей.** У исправного двигателя показатель регулярности гармоник (*harmonic regularity*) близок к единице (0,979), что указывает на устойчивую синусоидальную структуру. При несоосности регулярность несколько снижается (0,969), а при неисправности подшипника смещается акцент на показатель *harmonic ratio* (0,837), отражающий появление дополнительных частотных компонентов и возрастание гармонических искажений.

**4. Энергетический дисбаланс между фазами.** Во всех случаях неисправностей данный показатель выше, чем у здорового состояния (с 0,145 до 0,165–0,167), что говорит о нарушении симметрии нагрузок и изменении распределения мощности между фазами. Это типичный признак механических или электромагнитных дефектов, приводящих к неравномерному потреблению энергии.

**5. Спектральная корреляция фаз А–Б.** Наибольшая корреляция наблюдается у здорового двигателя (0,928), что указывает на синхронизированные токи фаз. При неисправности подшипника и при несоосности корреляция снижается до 0,910–0,911, отражая умеренное расхождение фазовых огибающих из-за гармонических возмущений. Системная несоосность даёт наименьшую корреляцию (0,903), что свидетельствует о более выраженном фазовом рассогласовании, вызванном взаимодействием двигателя и стенда (почему меньше чем у всех??).

**6. Доля мощности гармонических искажений (THD).** - При несоосности THD растёт (0,438), что указывает на появление выраженных гармонических искажений при сохранении умеренной синхронизации фаз. - Для подшипникового дефекта THD умеренный, а корреляция фаз близка к здоровому состоянию -> возникают отдельные гармоники от ударов, но общая синхронизация фаз сохраняется. - При системной несоосности THD также высока, но корреляция фаз минимальна -> гармонические искажения сопровождаются фазовым рассогласованием токов.

**7. Энтропия Винера (Spectral Flatness).** Для неисправного подшипника наблюдаются наименьшие значения ( $2,49 \times 10^{-5}$ ), что свидетельствует о спектре с выраженными гармоническими пиками. При несоосности стенда значение энтропии немного выше ( $2,91 \times 10^{-5}$ ), что указывает на более равномерное распределение спектральной энергии и шумоподобный характер огибающей.

**8. Коэффициент асимметрии (Skewness).** Значительные положительные значения и разброс коэффициента асимметрии ( $0,409 \pm 0,663$ ) характерны только для несоосности стенда, что говорит о несимметричных пульсациях момента и нестабильных колебательных процессах в системе «двигатель–нагрузка».

#### Обобщая результаты:

- СКЗ тока и признаки огибающей позволяют надёжно дифференцировать механические неисправности.
- Подшипниковая неисправность характеризуется ростом гармонических составляющих и снижением спектральной равномерности.
- Несоосность сопровождается увеличением RMS и THD при высокой межфазной корреляции сигналов.
- Несоосность стенда проявляется в снижении RMS (почему?????) и появлении статистических признаков неустойчивости (асимметрия (почему в несоосности не развивается???), рост энтропии).

Сравнение методов оценки важности показало, что метрики SHAP обеспечивают наиболее интерпретируемое распределение вклада признаков, в то время как Permutation Importance позволяет оценить их влияние на итоговую точность модели. Результаты всех методов взаимно подтверждают выделение ключевых признаков, что указывает на устойчивость полученной модели и корректность выбора признакового пространства.