Empirical probability and machine learning analysis of $m, n = 2, 1$ tearing mode onset parameter dependence in DIII-D H-mode scenarios

L. Bardóczi,[1, a)] N. J. Richner,[2, b)] J. Zhu,[3] C. Rea,[3] and N. C. Logan[4]
[1)] *General Atomics, P. O. Box 85608, San Diego, California 92186-5608, USA*
[2)] *Oak Ridge Associated Universities, Oak Ridge, Tennessee 37831, USA*
[3)] *Plasma Science and Fusion Center, MIT, Cambridge, MA, USA*
[4)] *Lawrence Livermore National Laboratory, Livermore, California 94551, USA*

(Dated: 20 June 2023)

$m, n = 2, 1$ tearing mode onset empirical probability and machine learning analyses of a multi-scenario DIII-D database of over 13,000 H-mode discharges show that the normalized plasma beta, the rotation profile and the magnetic equilibrium shape have the strongest impact on the 2,1 tearing mode stability, in qualitative agreement with neoclassical tearing modes (m and n are the poloidal and toroidal mode numbers, respectively). In addition, 2,1 tearing modes are most likely to destabilize when $n > 1$ tearing modes are already present in the core plasma. The covariance matrix of tearing sensitive plasma parameters takes a nearly block-diagonal form, with the blocks incorporating thermodynamic, current & safety factor profile, separatrix shape and plasma flow parameters, respectively. This suggests a number of paths to improved stability at fixed pressure and edge safety factor primarily by preserving a minimum of 1 kHz differential rotation, increasing the minimum safety factor above unity, using upper single null magnetic configuration and reducing the core impurity radiation. In addition, lower triangularity, lower elongation and lower pedestal pressure may also help to improve the stability. The electron and ion temperature, collisionality, resistivity, internal inductance and the parallel current gradient appear to only weakly correlate with the 2,1 tearing mode onsets in this database.

I. INTRODUCTION

Tearing modes are resistive magnetohydrodynamic (MHD) instabilities that relax the topology of the confining magnetic field of the tokamak plasma into an energetically more favorable configuration by forming magnetic islands on rational toroidal flux surfaces[1]. These islands are closed helical flux tubes, characterized by m poloidal and n toroidal mode numbers, whose helicity matches the local equilibrium magnetic field helicity. The $m, n = 2, 1$ magnetic islands (2, 1 in short) are the most common single root cause of sudden, uncontrolled and violent plasma terminations in tokamaks, called disruptions[2]. Due to their high disruptivity[3], 2,1 tearing avoidance and active stabilization are key tasks for scenario development of present day experiments and for operational regimes of future fusion reactors. This mission requires understanding the tearing onset mechanisms and their parameter dependence in reactor relevant conditions.

In plasmas characterized by sufficiently high bootstrap current (j_{BS}) fraction, the j_{BS} driven Neoclassical Tearing Modes (NTMs) can govern the evolution of the island once the island width (W) exceeds a threshold (W_{TR}). Their growth is facilitated by a hole in the pressure gradient driven j_{BS} in the center of the islands, which is caused by rapid transport on the island's internal nested flux surfaces. W_{TR} depends on the classical stability index (Δ'), the NTM drive[4–6] (which increases with the normalized plasma beta, β_N), the transport anisotropy within the island[7–11], the polarization current[12,13] and other small island terms that can be stabilizing or destabilizing. As there is no NTM drive without pressure profile flattening within the island, NTMs are non-linear instabilities and require another process to trigger the island growth. This trigger can be provided e.g. by a rapid core or edge MHD transient[13,14] (sawtooth crash or edge localized mode (ELM)), 3-wave coupling between multiple small islands[15] or the current profile if it is unstable[16]. Understanding the nature of tearing modes in reactor relevant plasmas is important, as it impacts decisions between experimentation paths in the development of fusion plasma scenarios. In order to project control solutions of tearing avoidance to future machines, a first principle based, causal and experimentally validated model is essential.

In this study, we experimentally characterize the sensitivity of 2,1 tearing onset to a group of local and global plasma parameters in a multi-scenario database of over half million time-slices in 13,000 high-confinement-mode (H-mode) DIII-D discharges and qualitatively compare the results to the modified Rutherford equation (MRE). This approach finds that the 2,1 magnetic island is dominantly pressure gradient driven, i.e. NTM. At fixed β_N, the instability onset rate is the most sensitive to the plasma differential rotation, followed by the magnetic equilibrium shape.

[a)] Electronic mail: bardoczil@fusion.gat.com
[b)] Electronic mail: richnern@fusion.gat.com

The rest of the paper is organized as follows. Section II describes the database and diagnostics. Section III describes the parameter dependence of the $m, n = 2, 1$ tearing mode onset probability. The relative importance of each parameter in the tearing onset is then investigated with machine learning analysis (ML) in Section IV. Finally, the summary and discussion of possible routes to improved stability are given in Section V.

II. DATABASE, DIAGNOSTICS AND ANALYSIS METHODS

In this section, we discuss the method of $m, n = 2, 1$ onset time determination (Subsection II A) , the analyzed database (Subsection II B) and the diagnostics (Subsection II C) used to infer the parameter dependence of 2,1 tearing mode onsets in DIII-D H-mode discharges.

A. Onset time determination

A multi-scenario database of 13,495 DIII-D H-mode discharges of a decade of experimental campaigns (2009 - 2019) are analyzed in this study. First, discharges that developed a large $n = 1$ rotating magnetic perturbation are selected, whose amplitude exceeds $\widetilde{B}_\theta = 7\,\mathrm{G}$ magnetic field signal at the tokamak wall. Next, we find an initial estimate for the 2,1 onset time (t_{onset}) in the unstable discharges by first finding the maximum of the $n = 1$ component of the $\widetilde{B}_\theta(t)$ signal and then working backward until $\widetilde{B}_\theta(t)$ hits a pre-defined threshold. Next, the (frequency-dependent) cross-power spectral density is calculated over incremental time windows (4 ms total time windows and 50% overlap) for two separate pairs of magnetic sensors [17]. The first pair is toroidally separated by $\phi = 32.7°$ angle in the low field side mid-plane. The second pair of probes is poloidally separated (in the low field side and high field side mid-plane) at the same toroidal angle. For each frequency bin and time window, the cross-phase is used to remove activity with $n \neq 1$ and m = odd. The coherence function (γ^2_{xy}) is used as an additional filter, removing any incoherent activity with $\gamma^2_{xy} < 0.7$. From this, the remaining power is dominated by coherent $m = 2, n = 1$ activity. The spectrogram of power vs. frequency and time is then used to estimate the onset time. First, the $P(f, t)$ point corresponding to the highest amplitude of filtered power is selected. Looking about a small frequency and time window, the algorithm iterates backward in time to points where filtered power is present. When no more power is found within the search window, the time coordinate of $P(f, t)$ is returned as the best estimator of the 2,1 onset time. Plasmas that did not develop rotating $2, 1$ islands or $\widetilde{B}_\theta < 7\,\mathrm{G}$, are considered stable in this study. Note that some of these discharges may still developed small rotating 2,1 islands, but those did not lead to disruptions.

B. Database

Within the selected database, there are 2,623 unstable and 10,872 stable plasmas where necessary diagnostics are available for the analyses described in this paper.

We construct a stable database and an unstable database for a set of local and global plasma parameters ($\{X^i\}$, defined in Section II C). We use both the stable and unstable discharges and we consider every $\Delta\tau \approx 100\,\mathrm{ms}$ window as an independent measurement point, as illustrated in Fig. 1. First, every stable plasma is subdivided into $\Delta\tau$ windows in the β_{N} flattop and time averages of X^i within each window are collected into the database. This procedure is repeated for the stable part of the unstable discharges, defined by the time stamps as $t < t_{\mathrm{onset}} - \Delta\tau$, and compiled into the stable database as well. Next, the unstable database of X^i is assembled from a single $\Delta\tau$ window preceding t_{onset} in each unstable discharge. This method yields an unstable database of up to $N_u = 2,645$ points and a stable database of up to $N_s = 509,535$ points for each X^i.

This method has the advantage of producing a large database for each X^i, allowing high resolution at low statistical noise. As not all relevant parameters within (and outside of) the X^i group are fixed when a particular X^j varies ($i \neq j$), correlations can exist between the X^i parameters. Such correlations can exist either due to physical relationships between the variables or due to preferred operational choices in the DIII-D experimental campaigns. We will examine such correlations in Section III.

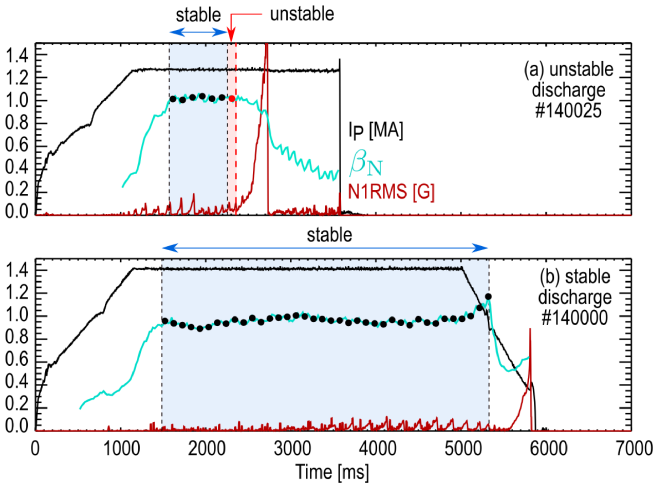

FIG. 1: Illustration of stable and unstable database assembly of a variable (β_{N} in this example). (a) Example unstable discharge and (b) example stable discharge. The β_{N} flattop of an unstable discharge is decomposed into a series of stable and one unstable time window. The β_{N} flattop of a stable discharge is decomposed into a series of stable time windows.

C. Diagnostics

With the goal to qualitatively test predictions of the modified Rutherford-equation, we assemble the stable and unstable databases of the $X^i = \{\beta_N, n_e, P_{ped}, T_i, T_e, P_{rad}, \nu_e^*, \nu_i^*, \eta, \kappa, \Delta, \square, \delta r_{sep}, A_2, \Delta f_{1,2}, \partial_\rho f_2, f_2, q_{min}, q_{95}, j_{BS}, r_1/a, r_2/a, \ell_i, \partial_r j_\parallel\}$ quantities:

- *Thermodynamics* — normalized plasma beta (β_N), line averaged electron density (n_e), electron and ion temperature at $q = 2$ (T_e and T_i, respectively), pedestal pressure (P_{ped}), electrical resistivity at $q = 2$ (η), electron and ion collisionality at $q = 2$ (ν_e^* and ν_i^*, respectively), core impurity radiation (P_{rad}, used as proxy for impurity concentration).
- *Magnetic equilibrium boundary shape* — elongation (κ), triangularity (Δ), squareness ($\square$) and radial distance to external 2nd separatrix (δr_{sep}). Note that $\delta r_{sep} > 0$ ($\delta r_{sep} < 0$) corresponds to upper (lower) single null magnetic configuration.
- *Plasma rotation* — rotation at $q = 2$ (f_2), rotation shear at $q = 2$ ($\partial_\rho f_2$) and differential rotation between $q = 1$ and $q = 2$ ($\Delta f_{1,2}$).
- *Current and safety factor profile parameters* — normalized internal inductance (ℓ_i), minimum safety factor (q_{min}), edge safety factor (q_{95}), bootstrap current at $q = 2$ (j_{BS}), parallel current gradient at $q = 2$ ($\partial_r j_\parallel$) and normalized minor radius of the $q = 1$ and $q = 2$ surfaces (r_1/a and r_2/a, respectively). Here, r_1 and r_2 are the minor radius coordinates of the $q = 1$ and $q = 2$ surfaces, respectively, and a is the plasma minor radius in the mid-plane.
- *Preceding tearing activity* — root-mean-square magnetic amplitude of rotating modes characterized by $n = 2$ toroidal mode number. This usually corresponds to the $m, n = 3, 2$ tearing mode (A_2).

β_N is calculated from the plasma β (kinetic pressure divided by magnetic pressure), minor radius (a), toroidal magnetic field (B_T) and plasma current I_p ($\beta_N = \beta a B_T / I_p$). The T_e and n_e profiles are measured via a combination of the Thomson scattering[18,19] and the Electron Cyclotron Emission[20,21] diagnostics. P_{ped} is calculated from a modified hyperbolic tangent fit to the edge temperature and density profiles. ν_e^*, ν_i^*, j_{BS} and η are calculated from the measured kinetic profiles via the AutoOneTwo automatic transport analysis code. T_i is the carbon impurity temperature and P is the tangential line of sight carbon impurity rotation, both measured by charge exchange recombination spectroscopy [22]. The local rotation shear at $q = 2$ is calculated as the derivative of rotation f with respect to the square root of the normalized toroidal flux surface label (ρ).

n_e is the line-averaged density calculated using the double-pass interferometer line-integrated density measurements and path lengths from magnetic equilibrium reconstructions [23].

P_{rad} is the core impurity radiation coming mostly from carbon, oxygen and any metals or other injected impurities present in the machine [24].

The plasma shape parameters, current and safety factor profiles are derived from the magnetic equilibrium reconstructed by the EFIT code[25], constrained by external magnetic probe data as well as internal measurements of the magnetic pitch angle by the motional Stark-effect diagnostic[26]. j_{BS} is calculated by the AutoOneTwo transport code and the parallel current gradient at $q = 2$ is provided by CAKE (Consistent Automatic Kinetic Equilibrium reconstruction[27]). Tearing stability dependence is explored via onset probability analysis in Section III and machine learning analysis Section IV.

III. PARAMETER DEPENDENCE OF 2,1 ONSET EMPIRICAL PROBABILITY

To characterize the sensitivity of the $m, n = 2, 1$ tearing onset to the $\{X^i\}$ parameter set, we calculate the onset empirical probability ($P(X^i)$, probability in short) with respect to each parameter.:

$$P(X^i) = \frac{H_u(X^i)}{H_t(X^i)} \tag{1}$$

Here $H_u(X^i)$ and $H_s(X^i)$ are the histograms of X^i calculated from the unstable and stable databases, respectively (hence the u and s indices), and $H_t(X^i) = H_u(X^i) + H_s(X^i)$ is the histogram of the union of the stable and unstable databases. All histograms are calculated in 35 points with respect to X^i. The ranges of the X^i are chosen to cover more than 95% of the H-mode operational space. The uncertainty of each point in H_u and H_s are calculated as $\sqrt{N}$, where N is the number of time-slices in a given bin. Gaussian error-propagation is then used to calculate the uncertainty of the total distribution and the uncertainty of the onset probability ($\delta P(X^i)$). Missing data is considered as *missing completely at random*[28] (MCR).

Note that the unit of $P(X^i)$ is arbitrary, as it depends on the length of the shot decomposition time window $\Delta\tau$ (described in Section II C), as there is always one unstable time slice in an unstable shot, regardless of the length of $\Delta\tau$, while the number of stable time slices are determined by the ratio of the stable operational time in the β_N flattop and $\Delta\tau$. Rescaling $P(X^i)$ by $\Delta\tau$ is not possible, as the choice of $\Delta\tau$ affects the length of the total stable & unstable time differently. As such, meaningful information is inferred by comparing the $P(X^i)$ empirical probabilities of different X^i parameters with each other, all calculated with the same $\Delta\tau$.

To assess the importance of each variable, we use the conservative estimate $\Delta P_i - \sigma_{P_i}$. Here σ_{P_i} is the maximum variation of $P(X^i)$ and σ_{P_i} is the associated uncertainty, which is calculated from the uncertainty of $P(X^i)$ via Gaussian error-propagation. We evaluate $\Delta P_i - \sigma_{P_i}$ in the ΔX^i interval that covers at least 95%

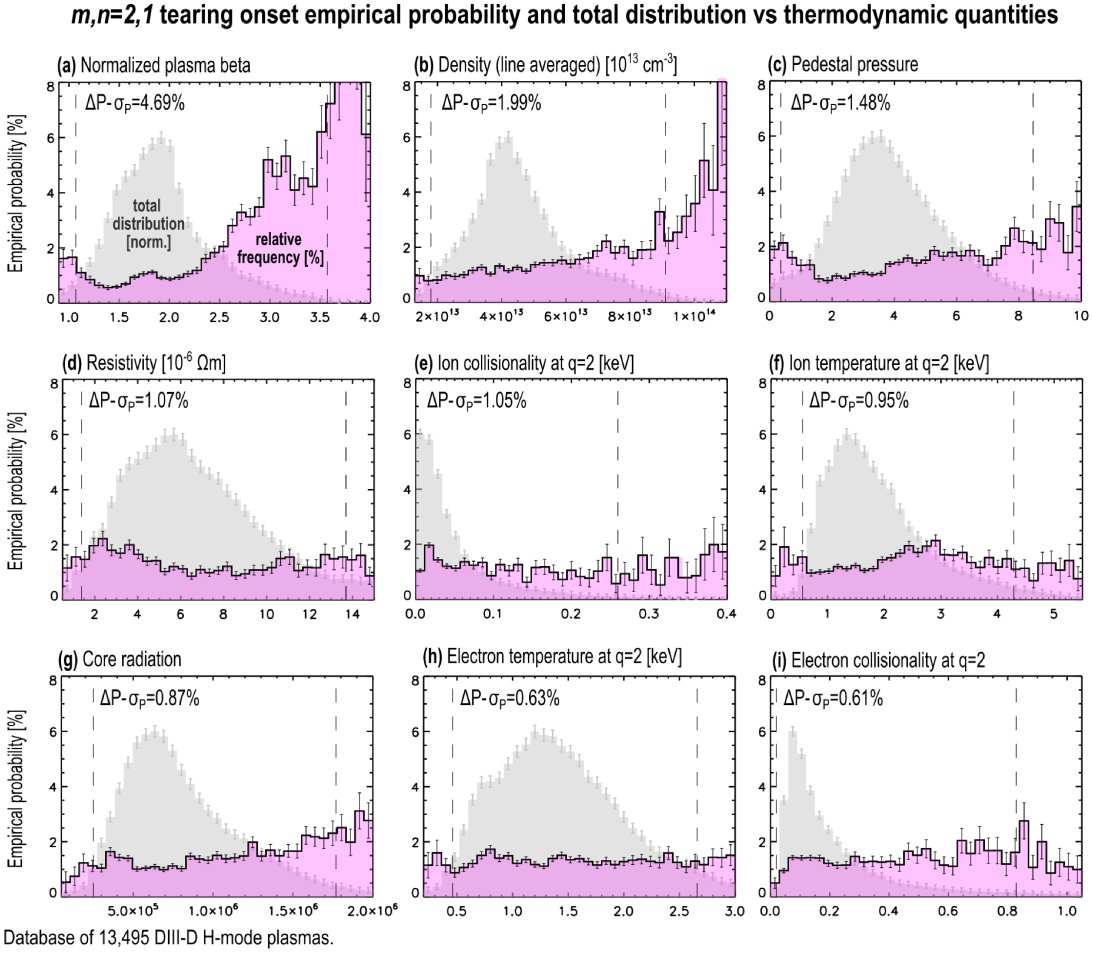

FIG. 2: Total distribution (gray) and empirical probability (colored). (a) Normalized plasma beta. (b) Line averaged electron density. (c) Pedestal pressure. (d) Resistivity at $q = 2$. (e) Ion collisionality at $q = 2$. (f) Ion temperature at $q = 2$. (g) Core impurity radiation. (h) Electron temperature at $q = 2$. (i) Electron collisionality at $q = 2$.

of the stable data points as well as 95% of the unstable data points. This ΔX^i interval is defined as the union of the intervals where $2.5\% < \mathrm{CDF}_s^i < 97.5\%$ and $2.5\% < \mathrm{CDF}_u^i < 97.5\%$ are both satisfied. CDF_s^i and CDF_u^i are the cumulative distribution functions of X^i in the stable and unstable databases, respectively. This range is shown with vertical dashed lines in Fig. 2, Fig. 3 and Fig. 4.

As mentioned earlier, correlations can exist between the examined parameters either due to physical relationships between the variables or due to preferred operational choices in the DIII-D experimental campaigns. In order to explore interdependence of the examined parameters in the database, we calculate the normalized covariance matrix:

$$C_{i,j} = \frac{\langle \widetilde{X}^i \widetilde{X}^j \rangle}{\sqrt{\langle (\widetilde{X}^i)^2 \rangle \langle (\widetilde{X}^j)^2 \rangle}} \quad \text{with} \tag{2}$$

$$\widetilde{X}^i = X^i - \langle X^i \rangle \tag{3}$$

where $\langle \cdots \rangle$ represents ensemble averaging over the stable database. The off-diagonal elements of $C_{i,j}$ are colored by their value in Fig. 5, while the diagonal elements are omitted. Interestingly $C_{i,j}$ is nearly block-diagonal as most correlations with large values occur within the thermodynamic block, within the shape parameter block, within the rotation parameter block and within the block of safety factor and current profile parameters. The strongest off-diagonal elements appear between the block of thermodynamic quantities and the block of current and safety factor profile parameters. Parameters within the magnetic equilibrium shape block, the rotation block and the preceding 3, 2 mode are weakly coupled to other parameters.

In the following, we discuss the importance of the X^i parameters using $P(X^i)$ and the $\Delta P_i - \sigma_{P_i}$ metric. Fig. 2 shows $P(X^i)$ with respect to thermodynamic quantities. Fig. 3 shows $P(X^i)$ with respect to plasma shape and plasma rotation quantities, as well as the amplitude of the preceding $m, n = 3, 2$ island. Fig. 4 shows $P(X^i)$ with respect to current and safety factor profile parameters. In each of these figures, the subfigures are ordered by $\Delta P_i - \sigma_{P_i}$. The colored graphs show $P(X^i)$, while $H_t(X^i)$ is shown as a gray shaded area in normalized units (the

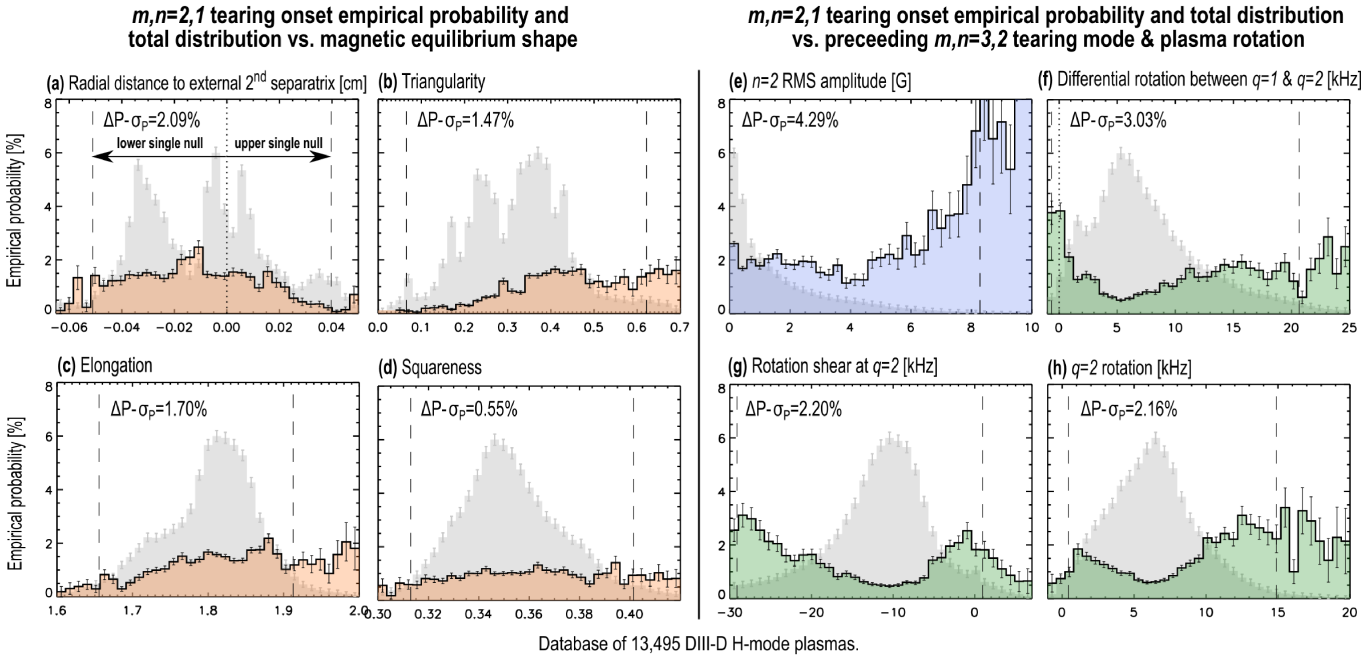

FIG. 3: Total distribution (gray) and empirical probability (colored) (a) Triangularity. (b) Radial distance to external second separatrix. (c) Elongation. (d) Squareness. (e) Differential rotation between $q = 1$ and $q = 2$, (f) $q = 2$ rotation (CER, ZipFit). (g) Rotation shear at $q = 2$ (CER, ZipFit). (h) Root-mean-square amplitude of the $m, n = 3, 2$ tearing mode.

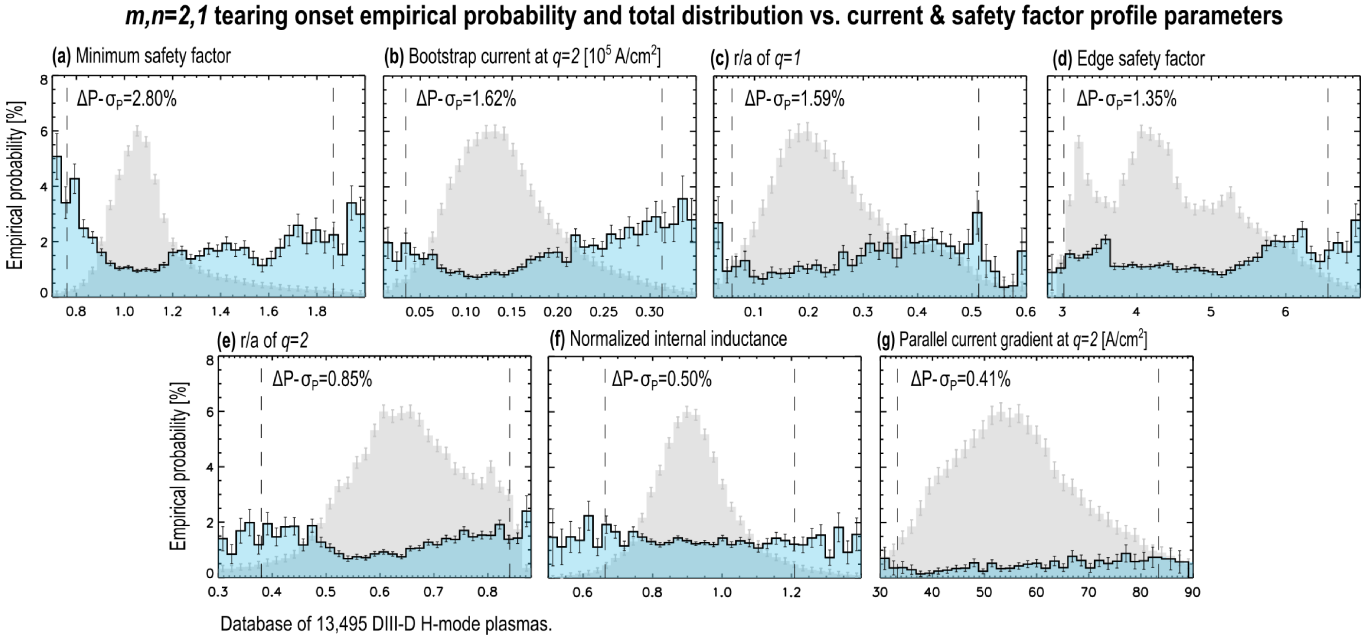

FIG. 4: Total distribution (gray) and empirical probability (colored). (a) Minimum safety factor. (b) Bootstrap current at $q = 2$. (c) Normalized minor radius coordinate of the $q = 1$ surface. (d) Edge safety factor. (e) Normalized minor radius coordinate of the $q = 2$ surface. (f) Normalized internal inductance. (g) Parallel current gradient at $q = 2$.

latter tells us how the machine was operated).

- *Thermodynamics* — $P(\beta_N)$ in Fig. 2 (a) shows that the mode onset probability strongly increases with β_N in the $\beta_N > 1.3$ operational space, which is known as a characteristic dependency of NTMs in the $\beta_N < 3.5$ range. Along with this, P increases with n_e [Fig. 2 (b)], P_{ped} [Fig. 2 (c)] and slightly with P_{rad} [Fig. 2 (g)]. Usu-

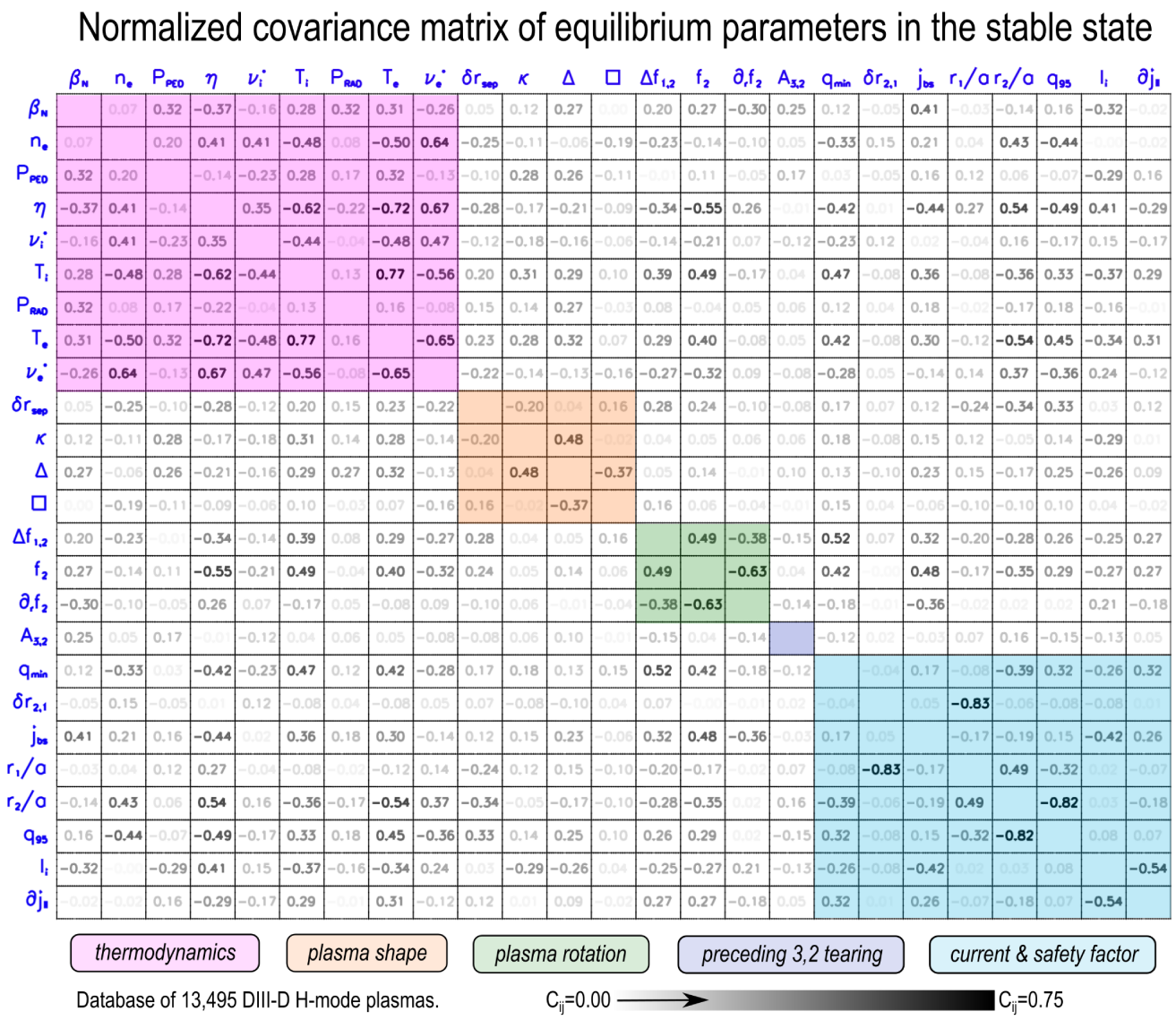

FIG. 5: Normalized covariance matrix of local and global parameters in the stable state. The diagonal elements are omitted.

ally, P is independent of η [Fig. 2 (d)], except at very low values of η. P slightly depends on T_i [Fig. 2 (f)], while nearly no dependence is seen with respect to ν_i^* [Fig. 2 (e)], T_e [Fig. 2 (h)] and ν_e^* [Fig. 2 (i)]. The different response to electrons and ions may be an indication of the role of polarization currents arising when the islands drift in the plasma frame[12,29].

Within this group of thermodynamic quantities, the strongest correlations exist between T_e and T_i, which indicates usually good collisional coupling between the electrons and the ions, strong anti-correlation between η and T_e as expected analytically[1], and consequently between T_i and η, see Fig. 5. T_e and T_i are both anti-correlated with n_e, which could at least in part be caused by the fact that most scenarios are operated around a target β_N while many experiments are aimed to test the effect of fueling and external heating. The collisionalities strongly depend on the temperatures and the density, as expected[1]. P_{rad} doesn't show strong correlation with either quantity. This is not surprising as the radiation strongly depends on the core impurity density, which is not directly correlated with the above quantities. Outside of the thermodynamic block these quantities mostly correlate with some quantities of the current and safety factor profile block: q_{min}, q_{95}, j_{BS} and naturally with the $q = 2$ surface location. The correlations with q_{min} and q_{95} is expected at least due to operational choices. For example consider that low q_{95} plasmas are often operated at lower β_N and vice versa, see for example the ITER baseline scenario (where $q_{95} \approx 3$, $q_{min} \approx 1$ and $\beta_N \approx 2$) and steady state advanced tokamak scenarios[30] (where $q_{95} > 4.5$, $q_{min} > 1$ and $\beta_N > 2$). The correlation with j_{BS} is also expected as the bootstrap current is proportional to the pressure gradient. The correlation with r_2/a reflects at least in part to the implicit dependence

(for example a change in the $q = 2$ location leads to a change of various thermodynamic quantities at the mode rational surface).

- *Magnetic equilibrium boundary shape* — The plasma shaping has considerable impact on the 2,1 tearing stability as well. P strongly increases with Δ [Fig. 3 (b)] and with κ [Fig. 3 (c)] and slightly with $\square$ [Fig. 3 (d)]. Within the spanned interval of κ and Δ the onset probability increases by about approximately a factor of nine. While Δ and κ are strongly correlated [Fig. 5] it is possible that their effect on the stability has somewhat different origins. The dependence on κ is in line with NTM theory as the bootstrap current fraction ($f_{\rm BS}$) increases with elongation[6] as $f_{\rm BS} \propto (1 + \kappa^2)$. On the other hand, drift kinetic theory[31] predicts that the NTM threshold increases with Δ, which is due to the modification of trapped particle orbits caused by Δ. Interestingly, P can improve by as much as a factor of about 8 when switching from lower single null to upper single null configuration (for example $P(\delta r_{\rm sep} \approx -4{\rm cm}) \approx 1.6$ but $P(\delta r_{\rm sep} \approx 4{\rm cm}) \approx 0.2$), see [Fig. 3 (a)] . $\delta r_{\rm sep}$ shows no significant correlation with other shape parameters, which suggests that the physics connecting $\delta r_{\rm sep}$ and P is independent of the effects caused by changes in κ and Δ, and so $\delta r_{\rm sep}$ may offer an independent degree of freedom to improve tearing stability. The shape parameter block is nearly independent of all the other examined quantities, with the exception of the correlation between the pedestal pressure and the triangularity.

- *Plasma rotation* — P shows significant variation with respect to $\Delta f_{1,2}$, $\partial_\rho f_2$ and f_2 [Fig. 4 (f), (g) and (h), respectively]. The plasma is most sensitive to the differential rotation in the narrow $|\Delta f_{1,2}| < 1\,{\rm kHz}$ range, showing that flat rotation profiles are prone to 2,1 onset. The rotation shear is optimal near $\partial_\rho f_2 \approx -8 - 12\,{\rm kHz}$, while both high and low $q = 2$ rotation are unfavorable, with an optimum near $f_2 \approx 5 - 7\,{\rm kHz}$, which is where most of the plasmas are operated. The correlation plot of $\Delta f_{1,2}$ and $\partial_\rho f_2$ (not shown) reveals that these quantities are correlated when both quantities are approaching zero and the correlation disappears at high differential rotation and high shear.

- *Preceding tearing activity* — Interestingly there is strong correlation between the onset of $2,1$ tearing modes and preceding $n = 2$ magnetic activity, which usually corresponds to $m, n = 3, 2$ islands. P increases by $> 4\%$ when large $n = 2$ modes are present in the plasma [Fig. 3 (e)]. This suggests, that preceding tearing modes either change the plasma conditions in an unfavorable way (for example by flattening the rotation profile[32]) or by directly participating in the $m, n = 2, 1$ island seeding (for example by three-wave coupling to the sawtooth precursor[15]). In either case, this suggests that A_2 can be used as an efficient 2,1 onset predictor. At low q_{95}, where the vast majority of the 2,1 islands disrupt, A_2 may be used for disruption prevention. If there is causality between A_2 and the 2,1 onset, then ECCD stabilization of the 3,2 island may also improve the success rate of disruption avoidance.

- *Current and safety factor profile* — $P(q_{\rm min})$ in Fig. 4 (a) shows that the plasmas are much more unstable when the $q = 1$ surface exists. This is consistent with the sawtooth instability triggering a significant fraction of the 2,1 islands. Clear dependence is observed on the bootstrap current [Fig. 4 (b)], in line with NTMs. The dependence on the $q = 1$ surface location [Fig. 4 (c)] may be due to the fact that at larger radii the temperature is lower, which makes the sawtooth less unstable. This reduces it's frequency which allows it to grow to larger amplitudes before the crash occurs. Some dependence on q_{95} is shown in Fig. 4 (d). This is not surprising as q_{95} is correlated with the distance between the $q = 2$ surface and the wall (2,1 islands in plasmas with higher q_{95} values sit further from the wall), as well as with $\beta_{\rm N}$ due to operational choices. The three peaks in the total distribution $H_t(q_{95})$ (gray) near $q_{95} = 3.1$, $q_{95} = 4.1$ and $q_{95} = 5.3$ represent the 3 preferred scenario groups that were run in DIII-D. P is higher both at low and high q_{95}, with a shallow minimum between about $q_{95} = 3.8$ and $q_{95} = 5.2$. On the other hand, $\partial_r j_\parallel$ is used here as a proxy to measure the possible role of the classical drive [Fig. 4 (g)]. The change of P within the entire range of $\partial_r j_\parallel$ is comparable to the uncertainty of the points and thus the effect of $\partial_r j_\parallel$ is weak. Further, $\ell_{\rm i}$ carries information about the current profile peaking [Fig. 2 (f)] whose evolution is correlated to the local current density at $q = 2$ as they are both governed by current diffusion on the resistive time-scale. If classical stability is affected significantly by the current relaxation, then a correlation between the mode onset and $\partial_r j_\parallel$ or $\ell_{\rm i}$ should be seen in the data. However, P has little to no dependence of the 2,1 island onset on these quantities.

The above results are summarized in Table II, where the parameters are ranked by $\Delta P_i - \sigma_{P_i}$. While lowering the density and bootstrap current are not reactor relevant ways of improve the stability, overall, there appear to be a number of independent, or partially independent paths to be explored. These analyses suggests that at given $\beta_{\rm N}$ the plasma stability can be improved by avoiding conditions when the $n = 2$ instability grows, avoiding flat rotation profiles, low minimum safety factors, running the plasmas in upper single null configuration is favorable, using magnetic configurations with the $q = 1$ surface deeper in the core, intermediate q_{95} and low impurity density are favorable. Importantly, it should be emphasized that these results are obtained by using all H-mode plasmas. The best strategy of tearing avoidance may vary between different scenario types. For example at low q_{95}, preceding $n > 1$ activity is clearly correlated with the appearance of 2,1 tearing modes[15]. On the contrary, the $m, n = 3, 2$ island prevents the 2,1 onset by keeping $q_{\rm min}$ above unity

Rank (1-12)	Parameter name	$\Delta P_i - \sigma_{P_i}$ [%]	Rank (13-24)	Parameter name	$\Delta P_i - \sigma_{P_i}$ [%]
1	β_{N}	4.69	13	Δ	1.47
2	A_2	4.29	14	q_{95}	1.35
3	$\Delta f_{1,2}$	3.03	15	η	1.07
4	q_{min}	2.80	16	ν_i^*	1.05
5	$\partial_r f_2$	2.20	17	$\mathrm{T_i}$	0.95
6	f_2	2.16	18	P_{rad}	0.87
7	δr_{sep}	2.09	19	r_2/a	0.85
8	$\mathrm{n_e}$	1.99	20	$\mathrm{T_e}$	0.63
9	κ	1.70	21	ν_e^*	0.61
10	j_{BS}	1.62	22	$\square$	0.55
11	r_1/a	1.59	23	ℓ_{i}	0.50
12	P_{ped}	1.48	24	$\partial_r j_{\parallel}$	0.41

TABLE I: Tested parameters ranked by their effect on the $m, n = 2, 1$ tearing onset probability (P).

in the hybrid scenario. Detailed investigation of different scenario types remains for future work.

In Section IV we investigate the importance of each parameter using machine learning analysis methods.

IV. MACHINE LEARNING ANALYSIS OF $m, n = 2, 1$ NTM ONSET PARAMETER DEPENDENCE

We rank the tearing sensitive parameters discussed in Section III using machine learning (ML) based linear and non-linear methods. The ML analyses require each included parameter to be measured at each time-slice. Hence those time slices must be removed from the input database where at least one diagnostic is unavailable. In order to maximize the number of input time-slices, and thereby maximize the fidelity of the ML analyses, we reduce the input parameter set to those that are the most important per the onset probability analysis presented in Section III and whose joint database reproduces the onset probability with the lowest χ^2. The included parameters are β_{N}, $\Delta f_{1,2}$, A_2, δr_{sep}, j_{BS}, $\mathrm{n_e}$, Δ, q_{min}, f_2, κ, P_{ped} and $\partial_r f_2$. We denote this reduced input parameter set with Z^i. These parameters are listed along with $\Delta P_i - \sigma_{P_i}$ and their corresponding ranks in Table II for convenience. In the following, we use the $\tilde{Z}^i = (Z^i - \langle Z^i \rangle)/\sigma_i$ transformed variables, where $\langle Z^i \rangle$) and σ_i are the ensemble average and standard deviation of $\tilde{Z}^i$, respectively. We then give the unstable time slices higher training weight (70) to balance the whole dataset as the ratio of stable samples to unstable samples is ≈ 130. This training weight is chosen to maximize the model performance.

A. Linear machine learning analysis

First, we employ the linear, statistical feature elimination method. We fit a logistic regression[33](LR) tearing mode prediction model using the full $\tilde{Z}^i$ set and evaluate the goodness of the fit using McFadden's pseudo R-squared statistic[34–36] (R_p^2). R_p^2 is an analogue of R^2 used in the ordinary least square regression. This statistic is a commonly used metric for evaluating LR models and it has good numerical and statistical properties[37]. R_p^2 has a range between 0 to 1 with higher values indicating better model fit. In practice, R_p^2 value falling between 0.2 and 0.4 indicates very good fit[37]. The obtained R_p^2 of the LR trained on the full $\tilde{Z}^i$ is our baseline value, which is 0.2088. Then, we remove one variable at a time (Z^k) to produce the $\tilde{Z}_k^i$ sub groups, each with N-1 variables (N is the number of variables in $\tilde{Z}_i$). The LR model is then refitted for each $\tilde{Z}_k^i$ set and R_p^2 is re-evaluated for each trained model. Finally, we compare the R_p^2 of each $\tilde{Z}_k^i$ set with the baseline value. If the obtained R_p^2 for one particular $\tilde{Z}_k^i$ list is much worse than the baseline value, then it means that removing the Z^k variable strongly affects the performance of the fitted model and so X^k is important for tearing mode onset prediction in this database. Therefore, a low R_p^2 value corresponds to high importance of a given Z^k parameter and R_p^2 gives the importance rank of Z^k.

The results of this analysis are summarized in Table II, which shows that the 2,1 instability is most sensitive to β_{N}, A_2 and q_{min}. These are consistent with the modes being pressure gradient driven, their onset preferentially occurs when $n > 1$ modes and sawtooth are present.

B. Non-linear machine learning analysis

To account for non-linear trends observed in Section III, next we use a non-linear ML model to analyze the tearing mode database. We fit an eXtreme Gradient Boosting (XGBoost) [38] prediction model, a highly efficient and flexible implementation of gradient boosted decision trees [39] on the training set and evaluate the performance of the model on the test set. Under a binary classification scheme, a successfully detected tearing mode onset is regarded as true positive (TP), while false positives (FP) correspond to a false warning, or a healthy plasma being declared to be tearing unstable. A trade-

	Onset rel. freq.		Linear ML		Non-linear ML		
Parameter	$\Delta P - \sigma_P$	**rank**	R_p^2	**rank**	**AUC**	**rank**	**Average rank**
β_{N}	4.69	1	0.1652	1	0.869	1	1.00
A_2	4.29	2	0.1925	2	0.882	4	2.67
$\Delta f_{1,2}$	3.03	3	0.2200	5	0.881	3	3.67
q_{min}	2.80	4	0.1959	3	0.890	9	5.34
δr_{sep}	2.09	7	0.2302	6	0.882	5	6.00
j_{BS}	1.62	11	0.2313	9	0.880	2	7.34
f_2	2.16	6	0.2319	10	0.884	7	7.67
κ	1.70	10	0.2132	4	0.890	10	8.00
$\mathrm{n_e}$	1.99	8	0.2324	11	0.883	6	8.34
Δ	1.47	14	0.2304	7	0.889	8	9.67
$\partial_r f_2$	2.20	5	0.2325	12	0.896	12	9.67
P_{ped}	1.48	13	0.2307	8	0.891	11	10.67

TABLE II: Summary of onset probability, linear and non-linear machine learning analysis results of DIII-D 2,1 tearing mode database. Parameters in this table are ordered by their average rank shown in the right-most column.

off can be achieved by adjusting the alarm threshold of the XGBoost model, as visually demonstrated by a receiver–operator characteristic (ROC) curve [40]. The area under the ROC curve (AUC) is used as performance metric for our XGBoost model. The obtained test AUC of the XGBoost model trained on all features is our baseline value. Then, we remove one feature at a time from the initial feature list and refit the XGBoost model with each reduced feature set and re-evaluate the test AUC for each trained model. Finally, we compare the test AUC of each case with the baseline value, which is 0.895. If the obtained test AUC for one particular feature list is much worse than the baseline value, it means that removing the chosen parameter strongly affects the performance of the fitted model which means it is a very important parameter for tearing mode onset prediction. Therefore, the low test AUC corresponds to high importance of the chosen signal and the test AUC of each case gives the importance rank of all considered signals.

The results are summarized in Table II. β_{N}, j_{BS} and $\Delta f_{1,2}$ rank the highest by AUC. In Table II we use the average rank as a combined metric based on the onset probability and machine learning analyses to order the tested parameters.

V. SUMMARY AND DISCUSSION

In this database study, we presented the $m, n = 2, 1$ tearing mode onset empirical probability with respect to the $\mathrm{X}^i = \{\beta_{\mathrm{N}}, \mathrm{n_e}, P_{\mathrm{ped}}, \mathrm{T_i}, \mathrm{T_e}, P_{\mathrm{rad}}, \nu_e^*, \nu_i^*, \eta, \kappa, \Delta, \square, \delta r_{\mathrm{sep}}, \mathrm{A}_2, \Delta f_{1,2}, \partial_\rho f_2, f_2, q_{\mathrm{min}}, q_{95}, j_{\mathrm{BS}}, r_1/a, r_2/a, \ell_{\mathrm{i}}, \partial_r j_\parallel\}$ parameter group using 512,180 time-slices of 13,495 DIII-D H-mode plasmas. The covariance matrix of these parameters was then analyzed to identify independent paths to improve the plasma stability. Finally, machine learning based linear and non-linear analyses were used to rank the parameters by their importance.

The combination of onset probability and machine learning analyses shows that the plasmas are more prone to tearing onset as β_{N} incresases. At a given β_{N}, 2,1 tearing modes are most likely to destabilize if $n > 1$ tearing modes are already present. In addition, the 2,1 onset greatly prefers near zero differential rotation and $q_{\mathrm{min}} < 1$. Such behavior is consistent with dominantly sawtooth seeding in concurrence with non-linear coupling to $n > 1$ modes. The mode onset preference of low $\Delta f_{1,2}$ suggests that the seed islands are stabilized by the differential rotation (e.g. through polarization currents). As such, $q_{\mathrm{min}} > 1$ and $\Delta f_{1,2} > 1\,\mathrm{kHz}$ are favorable for stability. The stability may further improve at intermediate rotation and rotation shear. Upper single null configuration, lower bootstrap current and $q = 1$ closer to the magnetic axis are favorable. Lower triangularity, elongation and pedestal pressure may also help to improve the stability. The electron and ion temperature, collisionality, resistivity, internal inductance and the parallel current gradient appear to only weakly correlate with the 2,1 tearing onset.

The covariance matrix of tearing sensitive parameters takes a nearly block-diagonal form, with the blocks incorporating thermodynamic, current & safety factor profile, separatrix shape and plasma flow parameters, respectively. This suggests a number of independent paths to improved stability at fixed pressure and edge safety factor and so a combination of parameter changes, as described above, may lead to substantial improvements.

The observed dependencies fit well into the neoclassical tearing mode theory. (i) The increasing onset probability with respect to β_{N}, P_{ped} and κ are consistent with bootstrap drive being at work. (ii) The increasing onset probability with respect to Δ is consistent with drift-kinetic effects increasing the NTM threshold island width[31]. (iii) In case of sawtooth triggered NTMs, the role of $\Delta f_{1,2}$ is consistent with stabilizing polarization currents becoming weaker when $\Delta f_{1,2}$ approaches zero, preceding the island onset. Correlated to this effect, the rotation shear also weakens.

The observed increase of tearing onset with impurity radiation in the core plasmas is consistent with radiation

driven magnetic islands[42]. Note that the gyroradius of impurities is much larger than the width of narrow islands near the NTM onset threshold and hence impurity trapping is unlikely to occur in narrow islands. However, even such a non-localized heat sink leads to a hollow electron temperature profile inside the islands if the electrons are collissionally coupled to cooling impurities[43]. Thus, it is possible that the observed radiation dependence is caused by radiation cooling decreasing the threshold island width.

The fact that the onset probability only slightly changes with the internal inductance and the parallel current gradient at $q = 2$ suggests that the classical stability parameter is nearly constant in the analyzed database.

The observed dependence on the single null location (this information is carried by δr_{sep}) remains an open question, especially considering that δr_{sep} shows no significant correlation with β_{N}, nor any other quantities considered in this study.

Finally, we emphasize that the presented results reflect to the average tearing mode onset behavior as they are derived from a combination of all types of H-mode plasmas in DIII-D. As such, the physics of the tearing modes and the parameter ranking may vary between different scenario types. Separate analyses of given scenarios are required in order to inform scenario specific strategies for tearing mode avoidance. Such analyses are left for future work.

ACKNOWLEDGEMENTS

The authors acknowledge very useful discussions with E. J. Strait, R. J. La Haye and J. D. Callen.

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Fusion Energy Sciences, using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Awards DE-FC02-04ER54698, DE-AC52-07NA27344 and DE-SC0014264. The data that support the findings of this study are available from the corresponding author upon reasonable request. **Disclaimer:** This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

REFERENCES

[1] J. Wesson. *Tokamaks*. Oxford University Press, 2011.
[2] P.C. de Vries, et al. *Nucl. Fusion*, 51(053018), 2011.
[3] R. Sweeney, et al. *Nucl. Fusion*, 57(016019), 2016.
[4] R. Carrera, et al. *Phys. Fluids*, 29:899–902, 1986.
[5] J.D. Callen, et al. *Plasma Physics and Controlled Nuclear Fusion Research 1986*, 2:157, 1986.
[6] A. G. Peeters. *Plasma Phys. Control. Fusion*, 42:B231–B243, 2000.
[7] R. Fitzpatrick. *Phys. Plasmas*, 2:825, 1995.
[8] S. Inagaki, et al. *Phys. Rev. Lett.*, 92(055002), 2004.
[9] J. A. Snape, et al. *Plasma Phys. Control. Fusion*, 54(085001), 2012.
[10] L. Bardóczi, et al. *Phys. Plasmas*, 23(052507), 2016.
[11] R Fitzpatrick. *Phys. Plasmas*, 24(052504), 2017.
[12] R. Fitzpatrick, et al. *Phys. Plasmas*, 13(122507), 2006.
[13] R.J. La Haye, et al. *Nucl. Fusion*, 62(056017), 2022.
[14] O. Sauter, et al. *Phys. Rev. Lett.*, 88(10), 2002.
[15] L. Bardóczi, et al. *Phys. Rev. Lett.*, 127(055002), 2021.
[16] P. H. Rutherdord. *Phys. Fluids*, 16:1903, 1973.
[17] E. J. Strait. *Rev. Sci. Instrum*, 77(023502), 2006.
[18] T. N. Carlstrom, et al. *Rev. Sci. Instrum*, 63(10):4901–4906, 1992.
[19] D. M. Ponce-Marquez, et al. *Rev. Sci. Instrum.*, 81(10D525), 2010.
[20] M. Bornatici, et al. *Nucl. Fusion*, 23(9), 1983.
[21] M. E. Austin et al. *Rev. Sci. Instrum.*, 74:1457, 2003.
[22] K. H. Burrell, et al. *Rev. Sci. Instrum*, 75(10):3455, 2004.
[23] M. A. Van Zeeland, et al. *Rev. Sci. Instrum*, 77(10F325), 2006.
[24] A. W. Leonard, et al. *Rev. Sci. Instrum.*, 66:1201–1204, 1995.
[25] L.L. Lao, et al. *Nuclear Fusion*, 30:1035, 1990.
[26] D. Wróblewski et al. *Rev. Sci. Instrum.*, 63:5140, 1992.
[27] Z.A. Xing, et al. *Fusion Engineering and Design*, 163(112163), 2021.
[28] Chapman & Hall. *Missing Data Analaysis in Practice*. Number 1482211920. CRC Press, 2015.
[29] J. D. Callen. Mhd transient seeding of disruptive neoclassical tearing modes. In *Bulletin of the American Physical Society*, volume 65. 2020.
[30] C M Greenfield, et al. *Plasma Phys. Control. Fusion*, 46(B213), 2004.
[31] A V Dudkovskaia, et al. *in preparation*, 2022.
[32] B. Tobias, et al. *Phys. Plasmas*, 23:056107, 2016.
[33] D. R. Cox. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2):215–232, 1958.
[34] P. D. Allison. Measures of fit for logistic regression. In NC: SAS Institute Inc Cary, editor, *In Proceedings of the SAS global forum 2014 conference*, pages 1–13. 2014.
[35] D. W. Hosmer Jr, et al. *Applied logistic regression*, volume 398. John Wiley and Sons, 2013.
[36] D. McFadden. *Conditional logit analysis of qualitative choice behavior*. Number 105-142. Academic Press, 1973.
[37] J. J. Louviere, et al. *Stated choice methods: analysis and applications*. Cambridge university press, 2000.
[38] T. Chen et al. A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 785–794. New York, NY, USA, 2016.
[39] J. H. Friedman. *Annals of statistics*, 29(5):1189–1232, 2001.
[40] A. P. Bradley. *Pattern Recognition*, 30(7):1145–1159, 1997.
[41] P. Emerson. *Social Choice and Welfare*, 40(2):353–358, 2016.
[42] D. A. Gates et al. *Phys. Rev. Lett.*, 108(165004), 2012.
[43] L. Bardóczi, et al. *Phys. Plasmas*, 24(062503), 2017.