Characterization and parametric dependencies of low wavenumber pedestal turbulence in the National Spherical Torus Experiment\textsuperscript{a)}

D. R. Smith,\textsuperscript{1,\textsuperscript{b)}} R. J. Fonck,\textsuperscript{1} G. R. McKee,\textsuperscript{1} D. S. Thompson,\textsuperscript{1} R. E. Bell,\textsuperscript{2} A. Diallo,\textsuperscript{2} W. Guttenfelder,\textsuperscript{2} S. M. Kaye,\textsuperscript{2} B. P. LeBlanc,\textsuperscript{2} and M. Podesta\textsuperscript{2}

\textsuperscript{1}Department of Engineering Physics, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA
\textsuperscript{2}Princeton Plasma Physics Laboratory, Princeton, New Jersey 08543, USA

(Received 29 November 2012; accepted 28 February 2013; published online 7 May 2013)

The spherical torus edge region is among the most challenging regimes for plasma turbulence simulations. Here, we measure the spatial and temporal properties of ion-scale turbulence in the steep gradient region of H-mode pedestals during edge localized mode-free, MHD quiescent periods in the National Spherical Torus Experiment. Poloidal correlation lengths are about 10 \( \rho_i \), and decorrelation times are about 5 \( a/c_s \). Next, we introduce a model aggregation technique to identify parametric dependencies among turbulence quantities and transport-relevant plasma parameters. The parametric dependencies show the most agreement with transport driven by trapped-electron mode, kinetic ballooning mode, and microtearing mode turbulence, and the least agreement with ion temperature gradient turbulence. In addition, the parametric dependencies are consistent with turbulence regulation by flow shear and the empirical relationship between wider pedestals and larger turbulent structures. © 2013 AIP Publishing LLC. [http://dx.doi.org/10.1063/1.4803913]

I. INTRODUCTION

Global confinement and first-wall heat load predictions in ITER and next-step devices depend on accurate models of the steep pedestal region. The spherical torus (ST)\textsuperscript{1} edge region is among the most challenging regimes for plasma turbulence simulations due to the inherent challenges of edge simulations and the distinct ST parameter regime with high \( \beta \) (2\( \mu_0 n_B/B^2 \)), large \( \rho_\ast (\rho_n/a) \), strong beam-driven flow, and strong shaping. Past results from the National Spherical Torus Experiment (NSTX)\textsuperscript{2} highlight novel turbulence and transport properties in ST plasmas. For instance, power balance analysis indicates electron thermal transport is the dominant loss mechanism, and ion thermal transport is at or near neoclassical values in NSTX beam-heated H-mode discharges.\textsuperscript{3,4} Stabilization or suppression of low-wavenumber (low-k) turbulence by strong equilibrium \( E \times B \) flow shear\textsuperscript{5} and field line curvature\textsuperscript{6} are leading explanations for near neoclassical ion thermal transport in NSTX beam-heated plasmas. Particle, momentum, and electron thermal transport remain anomalous and point to a turbulent transport mechanism. Also, power balance analysis indicates ion thermal transport decreases at higher plasma current, but the confinement time increase with plasma current in non-lithiated plasmas is weaker than that observed in conventional tokamaks.\textsuperscript{3,4,7} The high \( \beta \) regime makes ST plasmas more susceptible to low-k microtearing modes,\textsuperscript{8–10} and the scaling of NSTX confinement time with collisionality is consistent with collisional microtearing modes.\textsuperscript{11} Finally, recent turbulence measurements at the top of the H-mode pedestal during the ELM (edge localized mode) cycle were found to be consistent with ion-scale turbulence, such as ion temperature gradient (ITG), trapped electron mode (TEM), or kinetic ballooning mode (KBM) turbulence.\textsuperscript{12}

Edge and pedestal model validation motivates efforts to characterize low-k pedestal turbulence in the challenging ST parameter regime. Here, we characterize low-k pedestal turbulence quantities (\( k_0 \rho_i \approx 1.5 \), 0.8 < \( r/a < 0.95 \)) from beam emission spectroscopy (BES) measurements during ELM-free, MHD quiescent periods in NSTX H-mode discharges. In addition, we identify parametric dependencies among turbulence quantities and transport-relevant plasma parameters using a new model aggregation technique. Coherence spectra for poloidally adjacent channels exhibit broadband turbulence up to about 50 kHz. The turbulence parameters under investigation include poloidal correlation length, decorrelation time, and poloidal wavenumber. Poloidal correlation lengths in the pedestal are typically \( L_p \approx 15 \) cm and \( L_p/\rho_i \approx 10 \), and poloidal wavenumbers are typically \( k_0 \rho_i \approx 0.2 \). Also, decorrelation times are \( \tau_d/(a/c_s) \approx 5 \). The dimensionless quantities are similar to those observed in the core regions of L-mode tokamak discharges\textsuperscript{13} and consistent with drift-wave turbulence parameters. Next, a model aggregation algorithm identifies parametric dependencies among turbulence quantities and transport-relevant plasma parameters. Model aggregation is an analysis technique that identifies patterns in multi-dimensional datasets with complex interdependencies. Model aggregation can (1) identify more scalings than a single regression model and (2) produce a distribution of scaling coefficients covering a variety of model constraints. Observed scalings from model aggregation indicate \( L_p \) increases at higher \( \nabla n_e \), higher collisionality, and lower \( \nabla T_i \). Using heuristic transport models and turbulence theory, the observed scalings show the most agreement with transport driven by trapped-electron mode, kinetic ballooning mode, and microtearing mode turbulence, and the least agreement

\textsuperscript{b)}Invited speaker. Electronic address: drsmith@engr.wisc.edu
II. BEAM EMISSION SPECTROSCOPY OVERVIEW

The BES system on NSTX measures D$_e$ emission ($n = 3 \rightarrow 2$, $\lambda_0 = 656.1$ nm) from deuterium heating beams to study ion gyroscus fluctuations associated with low-k turbulence and instabilities. The beam velocity induces a Doppler shift in beam emission, and optical filters isolate beam D$_e$ emission from thermal D$_e$ emission. The intersection of optical sightlines and the neutral beam volume provides spatial localization, but a rigorous assessment of spatial and k-space properties requires point spread function calculations. BES measurements are sensitive to plasma density fluctuations with $\delta I_{D_e}/I_{D_e} = C (\delta n/n)$ where $C = C(E_{NB}, n, T_e, Z_{eff}) \approx 1/2$. $I_{D_e}$ is the beam D$_e$ emission intensity, and $n$ is the plasma density.

The NSTX BES system includes two optical assemblies centered at $R = 130$ and 140 cm ($r/a \approx 0.45$ and 0.85). The channel layout provides core-to-scrape-off layer radial coverage and four discrete poloidal arrays. Figure 1 shows fiber bundle images for the $R = 140$ cm optical assembly. BES measurements on NSTX are sensitive to fluctuations with $k_l \rho_s \leq 1.5$ where $\rho_s \approx 0.5 - 1.5$ cm is the ion sound gyroradius ($T_e, T_i \approx 0.3 - 1.0$). The optical views are aligned to typical NSTX pitch angles to optimize cross-field spatial resolution. New generation photodetectors and frequency-compensating, wideband preamplifiers provide photon-noise limited measurements at frequencies up to 400 kHz with refrigerant (non-cryogenic) cooling at $-20^\circ$C. The data acquisition system samples at 2 MHz, and digital anti-alias filters suppress high frequency thermal noise. The wideband, low noise detection system can also measure high frequency Alfvén and energetic particle modes up to about 800 kHz.

Figure 2 shows example BES measurements at two radial locations. BES signals show a sharp response at neutral beam injection (NBI) steps as expected, and BES auto-power spectra can exceed photodetector dark noise power spectra by 2–3 orders of magnitude. Notably, the emission d.c. amplitude in the core ($R = 129$ cm) exceeds edge emission ($R = 142$ cm) in Figure 2(f), but the edge emission shows higher amplitude turbulence at frequencies up to 100 kHz (Figure 2(g)). The larger signal-to-noise ratio at $R = 142$ cm and partially obscured signal changes at NBI steps is consistent with $(\delta n_e/n_e)_{edge} > (\delta n_e/n_e)_{core}$.

The BES channel layout provides multi-point measurements for assessing the spatial and temporal properties of turbulence. Poloidally separated channels in the poloidal array at $R = 140$ cm (Figure 1) provide measurements of the poloidal correlation length $L_p$, decorrelation time $\tau_d$, and poloidal wavenumber $k_\phi$ in the edge/pedestal region ($r/a \approx 0.80 - 0.95$). Figure 3 shows an example of turbulence quantities derived from poloidally separated measurements. Up to about 40 kHz, coherence decreases with poloidal separation and cross-phase increases with poloidal

![Image](336x227 to 420x297)

FIG. 1. NSTX cross section showing BES channels in the $R = 140$ cm poloidal array. Contour labels are normalized poloidal flux.

![Image](447x227 to 531x297)

FIG. 2. BES auto-power spectrograms for measurements at (a) $R = 129$ cm and (b) 142 cm; (c) low frequency odd-n magnetic fluctuations; (d) thermal D$_e$ emission measurements; (e) neutron measurements; (f) BES time-series data with NBI power (green line); and (g) BES auto-power spectra at 534 ms with photodetector dark noise spectrum.
separation as expected in Figures 3(a) and 3(b). The coherence exceeds the statistical noise floor and cross-phases are well-resolved. Similarly, auto-power spectra show BES signals exceed dark noise and photon noise levels up to about 40 kHz in Figure 3(c). Filtered data (8–50 kHz) in Figure 3(d) show the poloidal motion of turbulence structures from the reference channel to the maximum poloidal separation, \( \Delta z = 6 \) cm. Turbulence quantities such as correlation length are calculated from time-lag auto- and cross-correlation functions in Figure 3(e). 15–40 ms data windows are segmented into 30–150 bins with 512–2048 time points per bin. Coherence, cross-phase, and power spectra and time-lag correlation functions in Figure 3 are bin-averaged quantities. The cross-correlation envelope (calculated using the Hilbert transform) at zero time-lag decreases with channel separation, and the poloidal correlation length \( L_p \) is the corresponding 1/e length as shown in Figure 3(f). The time-lag of the peak cross-correlation envelope increases with channel separation as expected, and the linear relationship gives the eddy poloidal group velocity \( v_g \) as shown in Figure 3(h). The peak cross-correlation envelope decreases with time-lag, and the decorrelation time \( \tau_d \) is the corresponding 1/e time-lag as shown in Figure 3(g). Note that the decorrelation time calculation does not require transforming to the plasma frame; transforming to a frame moving in the plasma direction would slide data points in Figure 3(g) down the curve in a manner that preserves the decorrelation time. In essence, the poloidally separated measurements function as a fixed measurement in the plasma rest frame. The poloidal wavenumber \( k_h \) is derived from the eddy size inferred from the eddy velocity and time-lag between auto-correlation anti-nodes as shown in Figure 3(e).

Measured backlit fiber images are about 3 cm at the center of the neutral beam, and optical modeling indicates fiber images are 3.75 cm across at full-width half-max (FWHM, 50% of peak intensity). Accurate spatial and k-space characterization of BES measurements require PSF calculations that convolve optical system properties with image distortion from neutral beam profiles, magnetic field geometry, and atomic excited state lifetimes. Recent full-physics PSF calculations indicate image distortion at all locations is generally mild with FWHM image sizes in the range 3.6–3.9 cm. In the low density edge region at \( R \approx 140 \) cm, atomic excited state lifetimes are the dominant contribution to radial image distortion. In the core at \( R \approx 133 \) cm, magnetic field misalignment is the dominant contribution to image distortion (radial and poloidal). Accordingly, the measurements are sensitive to fluctuations with \( k \approx 2.5 \) cm\(^{-1} \) and \( k_{p*} \approx 1.5 \). PSF corrections are not applied to measurements below because the mild image distortion in the edge region is radial, not poloidal.

III. PEDESTAL TURBULENCE MEASUREMENTS IN ELM-FREE, MHD QUIESCENT H-MODE PLASMAS

ELM-free and MHD quiescent periods in NSTX H-mode discharges were identified and data from four or five poloidally separated BES channels at \( R = 140 \) cm (Figure 1) were analyzed to study low-k pedestal turbulence with \( k_o \rho_i \approx 0.2 \) and \( 0.8 < r/a < 0.95 \). Characteristic discharges and times of interest are listed in Table I. BES measurements provided poloidal correlation length, poloidal wavenumber, and decorrelation time (Figures 3(e)–3(g) in the pedestal region. BES signals were frequency filtered to isolate 8–50 kHz components, the typical frequency range for observed broadband turbulence. The 8 kHz lower limit ensures low frequency beam oscillations (\( f \approx 5 \) kHz) do not

<table>
<thead>
<tr>
<th>Discharge</th>
<th>Time of interest</th>
</tr>
</thead>
<tbody>
<tr>
<td>141125</td>
<td>925–950 ms</td>
</tr>
<tr>
<td>141131</td>
<td>700–725 ms</td>
</tr>
<tr>
<td>141147</td>
<td>730–750 ms</td>
</tr>
<tr>
<td>141255</td>
<td>740–760 ms</td>
</tr>
</tbody>
</table>
FIG. 4. (a)-(d) Profile measurements from the database and BES measurement location (grey box), (e) inferred $E_r$ profiles from charge exchange measurements (neglecting poloidal rotation), (f) distributions of poloidal correlation lengths, (g) decorrelation times, and (h) poloidal wavenumbers in the database.

contaminate turbulence signals, though power spectra (Figure 3(c)) show no evidence of beam oscillations. Discharges with $B_T = 4.5$ kG, $I_p = 700–900$ kA, and lower single-null geometry were screened for ELM-free, MHD quiescent periods that persisted at least 200 ms. Plasma parameters slowly evolved during ELM-free, MHD quiescent periods, so long time windows were partitioned into shorter windows. In total, 129 times of interest with 15–45 ms duration were identified from 29 discharges. Turbulence quantities and plasma parameters were averaged over each time window. Figure 4 shows plasma profiles and distributions of turbulence quantities in the database. Multi-point Thomson scattering provides electron density and temperature ($n_e$ and $T_e$) measurements, and charge exchange spectroscopy provides ion temperature and toroidal velocity measurements ($T_i$ and $V_t$).21 Radial electric field ($E_r$) profiles are inferred from carbon density, temperature, and toroidal velocity. The poloidal velocity contribution to $E_r$ is neglected because past results suggest the poloidal velocity contribution is small even in the pedestal.22 BES measurements in the steep gradient region show poloidal correlation lengths are $L_p \approx 10–20$ cm, poloidal wavenumbers are $k_\theta \approx 0.1–0.2$ cm$^{-1}$, and decorrelation times are $\tau_d \approx 10–20$ ms.

Table II lists 10th and 90th percentile ranges for turbulence quantities and plasma parameters in the database. Dimensionless turbulence quantities satisfy $L_p/\rho_i \sim 10$, $k_\theta \rho_i \sim 0.2$, and $\tau_d/(a/c_i) \sim 5$. Poloidal correlation lengths in Table II for low-field NSTX plasmas are generally longer than previously reported correlation lengths in high-field tokamak plasmas,13 but dimensionless parameters like $L_p/\rho_i$ and $\tau_d/(a/c_i)$ are similar. The quantity $k_\theta L_p \approx 2$ can be understood in terms of Gaussian distributions in real-space and k-space. For example, a spatial distribution $\exp(-(x-x_0)^2/\Delta x^2)$ produces a k-space distribution $\exp(-(k-k_0)^2/\Delta k^2)$ with $\Delta x\Delta k = 2$. With $k_\theta L_p \approx 2$ and $\Delta x \approx L_p$, it is easy to show $k_\theta/\Delta k \approx 1$, which represents a Gaussian distribution with similar width and shift. Also, in Table II, the equilibrium toroidal flow shear generally exceeds the turbulence decorrelation rate ($\nabla V_t > \tau_d^{-1}$), a puzzling result. However, $E \times B$ shear rates are $\omega_{E \times B} \approx 50–200$ kHz, so $\omega_{E \times B} \approx \tau_d^{-1}$ as expected, $\omega_{E \times B}$ values are not tabulated in Table II because inferred values are susceptible to large errors associated with second derivatives of pressure profiles. Plasma parameters in Table II generally show $50$–$300$% variation except for inverse aspect ratio $\epsilon$, elongation $\kappa$, lower triangularity $\delta_1$, $\rho_\theta^*$, and $\rho_\|^*$. The lack of variation in $\epsilon$, $\kappa$, and $\delta_1$ is likely due to screening for ELM-free, MHD-quiescent H-mode discharges. In the following sections, we untangle parametric dependencies among pedestal turbulence quantities and plasma parameters in Table II, but $\epsilon$, $\kappa$, $\delta_1$, $\rho_\theta^*$, and $\rho_\|^*$ are omitted from analysis due to lack of variation. Note that the nonlocal edge parameters such as pedestal width, height, and separatrix separation ($\Delta R_{ped}$, $n_{ped}$, and $\delta_{sep}^*$, respectively) are included in the database due to their possible impact on pedestal turbulence. Finally, radial correlation lengths ($L_r$) from the BES radial array (Figure 1) are not tabulated in Table II because preliminary analysis indicates $L_r$ in the steep gradient region is less than the BES image spot size, so the analysis only provides an upper bound with $L_r \lesssim 3$ cm. However, the upper bound on $L_r$ provides a lower bound for eddy anisotropy, $L_p/L_r \geq 3$. For comparison, pedestal turbulence measurements in DIII-D tokamak plasmas showed smaller anisotropy with $L_p/L_r \approx 1.2–1.5$.40

IV. REGRESSION ANALYSIS AND MODEL AGGREGATION

With a database of measured turbulence quantities and plasma parameters in hand, we now identify parametric dependencies using a stepwise multivariate linear regression (SMLR) algorithm and model aggregation. Let $y_i$ denote turbulence quantities such as correlation length, and let $x_{i,j}$ denote plasma parameters such as density gradient ($i$ indexes database entry and $k$ indexes plasma parameter). The SMLR algorithm finds models in the form

$$\text{SMLR}: y_i = \beta_i + \sum_{j=1}^{k} \beta_{i,j} \cdot x_{i,j} + \epsilon_i,$$

where $\beta_i$ is the intercept and $\beta_{i,j}$ are the regression coefficients. The regression analysis is performed using the R statistical software package. The data are divided into training and test sets, and the model is trained on the training data. The performance of the model is assessed on the test data, and the model is refined iteratively until a satisfactory fit is obtained. The final model is then used to predict turbulence quantities from plasma parameters. The model aggregation step involves combining multiple SMLR models to improve the accuracy of the predictions. This is achieved by selecting a subset of the most significant features from each model and combining them using a weighted average approach. The resulting model is then used to make predictions on new data. The effectiveness of this approach is evaluated by comparing the predictions with the measured data, and the model is refined further if necessary. The final model is then used to make predictions on new data.
can be highly subjective due to numerous statistical metrics and problematic due to potential parameter preferences. Previous turbulence scaling results scanned a single dimensionless parameter, such as \( \rho^2 \), while holding other transport-relevant parameters fixed.\textsuperscript{15-23} Here, we introduce and implement a model aggregation technique to identify parametric dependencies among turbulence quantities and transport-relevant plasma parameters. The combination of SMLR and model aggregation is an exploratory technique to identify patterns in multi-dimensional datasets with complex interdependencies. Other exploratory data techniques include maximal information-based nonparametric exploration,\textsuperscript{26} distance correlation,\textsuperscript{27} and hierarchical clustering.\textsuperscript{28} Model aggregation can be considered a “model of models” or a type of meta-analysis. Model aggregation produces \( z_k \) distributions from models identified by the SMLR algorithm. To illustrate the advantage of model aggregation, consider the six regression models for \( L_p/\rho_i \) in Table III. The individual models in Table III provide parametric scalings for three or four plasma parameters with other parameters unconstrained. In aggregate, the models provide multiple values of \( z_k \) coefficients for all plasma parameters under a variety of constraints. The emergence of consistent scalings from multiple models with a variety of constraints boosts confidence in the scalings. In summary, model aggregation provides (1) \( z_k \) scaling coefficients for more plasma parameters than a single model and (2) a distribution of \( z_k \) coefficients covering a variety of constraints.

Models identified by the SMLR algorithm are screened for multicollinearity and residual normality to ensure statistical properties indicative of valid regression models. Multicollinearity is the linear dependence among regression variables (\( x_k \)), and excessive multicollinearity inflates the uncertainty of \( z_k \) coefficients.\textsuperscript{24} Non-normal residual distributions violate the mathematical framework of regression analysis. Table IV summarizes the models identified by the

\[
\hat{y}_j - \bar{y} = \sum k \alpha_k x_{kj} - \bar{x}_k, \tag{1}
\]

where \( \sigma \) are standard deviations for \( y_j \) and \( x_{kj} \) and \( \bar{y} \) are turbulence quantities predicted by the model. The \( \alpha_k \) coefficients are the linear scaling coefficients when other plasma parameters in the model are fixed; parameters absent from the model are unconstrained. Also, the \( \alpha_k \) coefficients are dimensionless and directly comparable due to the normalization in Eq. (1). The SMLR algorithm minimizes the model’s squared sum of errors, \( \text{SSE} \equiv \sum (y_j - \hat{y}_j)^2 \), by adding or removing \( x_k \) parameters such that the inferred significance of each \( \alpha_k \) value exceeds 95\%.\textsuperscript{24} More technically, the inferred significance of each \( \alpha_k \) value exceeds 95\% when the probability of the null hypothesis (\( H_0 : \alpha_k = 0 \)) is less than 5\% according to the \( t \)-statistic associated with \( \alpha_k \).\textsuperscript{24}

The SMLR algorithm searches the high dimensional \( x_k \)-space for regression models at SSE local minima. Many SSE local minima can exist, so the SMLR algorithm can identify numerous regression models by starting from different initial states. A single regression model provides a limited set of \( \alpha_k \) scaling coefficients that are applicable only when other parameters in the model are fixed. In addition, selecting the “best” regression model from candidate models

\begin{table}[h]
\centering
\caption{Database quantities.}
\begin{tabular}{llll}
\hline Parameter & Range\textsuperscript{a} & Parameter & Range\textsuperscript{a} \\
\hline Turbulence quantities & & & \\
\hline \( L_p \) (cm) & 9.5–19 & \( k_q L_p \) & 1.2–2.8 \\
\( L_p/\rho_i \) & 7.6–18 & \( \tau_x (\mu s) \) & 8.6–28 \\
\( L_p/\rho_i \) & 9.0–21 & \( \tau_x^s (\mu s) \) & 4.6–37 \\
\( k_q \) (cm\textsuperscript{-1}) & 0.07–0.25 & \( \tau_x^c/\sigma_x \) & 2.8–22 \\
\( k_q \) (cm\textsuperscript{-1}) & 0.07–0.31 & \( \tau_x^c/\sigma_x \) & 1.1–8.6 \\
\( k_q \) (cm\textsuperscript{-1}) & 0.06–0.25 & \( \tau_x/(\sigma_x c_i) \) & 2.6–7.6 \\
\hline Plasma parameters & & & \\
\hline \( n_p (10^{13}/\text{cm}^3) \) & 1.7–2.6 & \( \rho_i^1/\rho_i^2 \) & 0.017–0.021 \\
\( \nabla n_p (10^{13}/\text{cm}^3) \) & 0.56–0.90 & \( \rho_i^1/\rho_i^2 \) & 0.021–0.026 \\
1/\( \text{L}_{pe} \) (cm\textsuperscript{-1})\textsuperscript{b} & 0.28–0.44 & \( \sigma_x^2 \) & 0.78–0.52 \\
\( T_e \) (keV) & 0.11–0.19 & \( q \) & 5.9–9.7 \\
\( \nabla T_e \) (keV/cm) & 0.061–0.094 & \( \delta \) & 2.5–5.5 \\
1/\( \text{L}_{Te} \) (cm\textsuperscript{-1})\textsuperscript{b} & 0.47–0.64 & \( \epsilon \) & 0.56–0.63 \\
\( \text{L}_{Te} \) (keV) & 0.33–0.50 & \( \kappa \) & 2.4–2.5 \\
\( \nabla \text{L}_{Te} \) (keV/cm) & 0.03–0.15 & \( \delta_i \) & 0.61–0.73 \\
1/\( \text{L}_{Te} \) (cm\textsuperscript{-1})\textsuperscript{b} & 0.07–0.34 & \( \nu_r (10^7/s) \) & 0.43–0.80 \\
\( V_e \) (km/s) & 37–68 & \( \nu_e (10^7/s) \) & 0.78–0.80 \\
\( \nabla V_e \) (10^8/s) & 0.33–1.7 & \( \nu_r (10^7/s) \) & 1.5–3.5 \\
\( E_r \) (V/cm) & 9.7–100 & \( \nu_r (10^7/s) \) & 0.70–0.21 \\
\( n_{ped} (10^{13}/\text{cm}^3)\textsuperscript{a} \) & 5.9–8.1 & \( \nu_r \) & 3.0%–5.3% \\
\( \Delta R_{ped} \) (cm) & 15–22 & \( \nu_r \) & 0.69%–1.6% \\
\hline
\end{tabular}
\textsuperscript{a}10th–90th percentile range.
\textsuperscript{b}1/Lx \equiv \nabla x/X.
\textsuperscript{c}(\sigma_x \sigma_r)/(\sigma_r/\epsilon_B).
\textsuperscript{d}1/\text{L}_{Te} \equiv 2\rho_i/\rho_i^2 \times B_\perp^2/\beta_r.
\textsuperscript{e}Pedestal height \( n_{ped} \) and width \( \Delta R_{ped} \) from electron density profile piecewise fits to linear and tanh functions with continuous first derivative.
\textsuperscript{f}Outboard radial distance to second separatrix; \( \sigma^\mu \), \( \sigma^\nu \) for lower single null configuration.
\end{table}

\begin{table}[h]
\centering
\caption{\( z_k \) and cross-correlation (\( C_{jk} \) values for a subset of \( L_p/\rho_i \) models. Parentheses around \( C_{jk} \) values indicate the \( z_j \)–\( z_k \) parameter pair is prohibited in models due to large cross-correlation.}
\begin{tabular}{cccccccc}
\hline Parameter & \( \nabla n_p \) & \( T_e \) & \( T_i \) & 1/\( \text{L}_{Te} \) & \( \nabla V_e \) & \( \nu_r \) & \( n_{ped} \) \\
\hline Model \( R^2 \) & & & & & & & \\
0.63 & 0.28 & … & -0.20 & -0.29 & 0.31 & … & … \\
0.63 & 0.34 & … & … & -0.37 & 0.30 & … & … \\
0.61 & 0.46 & -0.21 & … & -0.38 & … & … & … \\
0.60 & … & … & … & -0.47 & 0.38 & 0.24 & … \\
0.60 & … & … & -0.22 & -0.35 & 0.40 & 0.15 & … \\
0.55 & … & -0.24 & … & -0.55 & 0.36 & … & … \\
\hline \end{tabular}
\begin{tabular}{cccccccc}
\hline Parameter & \( \nabla n_p \) & \( T_e \) & \( T_i \) & 1/\( \text{L}_{Te} \) & \( \nabla V_e \) & \( \nu_r \) & \( n_{ped} \) \\
\hline Model \( C_{jk} \) values & & & & & & & \\
0.74 & -0.12 & -0.04 & -0.14 & 0.07 & 0.38 & (1.0) \\
0.59 & -0.83 & 0.27 & (0.62) & 0.63 & (1.0) \\
-0.33 & 0.27 & (0.62) & 0.63 & (1.0) \\
-0.38 & 0.08 & 0.28 & (1.0) \\
-0.32 & 0.44 & (1.0) \\
-0.26 & (1.0) \\
\hline
\end{tabular}
\end{table}
SMLR algorithm and lists several statistical quantities that characterize multicollinearity and residual normality. The SMLR algorithm was initialized with about 6000 parameter combinations, and the algorithm returned 12–50 unique models for each turbulence quantity. An example SMLR calculation for $k_0p_{\rho_s}$ starts with $V_i$ and $\nu_i$ in the initial state. Next, the algorithm adds $T_i/T_j$, then $\nabla n_i$, then $\nabla V_i$, then $\beta_j$, and then removes $T_j/T_i$ and $\nabla n_i$. As the algorithm converges on a final model, $R^2$ values (coefficient of determination, or goodness of fit; $R^2 = \sum(y_i - \bar{y})^2/\sum(y_i - \bar{y})^2$) increases from 0.59 to 0.67. $R^2$ values in Table IV indicate the models generally captured 30%–70% of the variation in the turbulence quantities. Also, about 100–150 initial states converged on redundant final models for each turbulence quantity. For the initial states $\{T_i, \nabla T_i\}, \{\nabla n_i, T_i\}$, and $\{n_i, T_i\}$ converged on the same final model $\{n_i, \nabla n_i, T_i, \nabla T_i\}$ for $L_p$, but the initial state $\{\nabla n_i, T_i\}$ converged on a different final model. Finally, the quantities $\tau_d \omega_{pe}^* \tau_d \omega_{pe}^*$, and $\tau_d \omega_{pe}^*$ are absent from Table IV because the SMLR algorithm failed to identify models for those quantities. The remainder of this section describes the statistical tests for multicollinearity and residual normality, and Sec. V describes the parametric scalings that emerge from model aggregation.

Two strategies are employed to screen for excessive multicollinearity in regression models. First, plasma parameter pairs with large cross-correlation are prohibited from models. The cross-correlation for parameters $x_i$ and $x_j$ is $C_{ij} = \sum (x_i - \bar{x}_i)(x_j - \bar{x}_j)/\sigma_i \sigma_j$. $C_{ih}$ is specified in the SMLR algorithm, and parameter pairs with $|C_{jk}| > C_{ih}$ are prohibited from models. The example in Table III lists $C_{jk}$ values for all parameter pairs, and parameter pairs with $|C_{jk}| > 0.6$ are denoted by parentheses and prohibited from models. The SMLR algorithm was run with $C_{ih} = 0.5 - 0.8$ to verify results are consistent across a range of correlation thresholds. Table IV summarizes all models identified by the SMLR algorithm. The $C_{ih}$ limits in Table IV correspond to 0.6 for correlation length and wavenumber models and 0.8 for decorrelation time models, and the associated scalings and models are consistent with other $C_{ih}$ values. Next, multicollinearity among three or more $x_i$ parameters is assessed using the variance inflation factor VIF$_k$ for each parameter $x_i$ in a model ($VIF_i \equiv 1/(1 - R_i^2)$ where $R_i^2$ is the squared multiple correlation that quantities the variability in $x_i$ captured by variation in other $x_j/k$). Models with VIF$_k > 10$ be susceptible to large uncertainties in $a_k$ parameters.

Table IV lists the range of max(VIF$_k$) values for models identified by the SMLR algorithm. Correlation length and wavenumber models exhibited max(VIF$_k$) < 4, and decorrelation time models exhibited max(VIF$_k$) < 7. Accordingly, max(|$C_{jk}$| and max(VIF$_k$)) values in Table IV are sufficiently low and indicate multicollinearity is not excessive.

Residual normality includes the independent and normal distribution of residuals $(r_i \equiv y_i - \bar{y}_i)$ and the absence of residual outliers. Non-normal residual distributions violate the mathematical framework of regression analysis. To assess residual normality, residuals are screened for outliers and the skewness and kurtosis of residual distributions are calculated. Studentized residuals, $r_i/\sqrt{\text{var}(r_i)}$, follow a t-distribution, and models with max(|$r_i/\text{ib}$| > 1, where $\text{ib}$ is the 95% significance level for the t-distribution, may contain outliers that distort the regression model. Correlation length, wavenumber, and decorrelation time models in Table IV exhibit max(|$r_i/\text{ib}$| < 1, so model distortion by outliers is unlikely. Skewness $\text{Sk} \equiv E(r_i - \bar{r})^3/\sigma_r^3$ and excess kurtosis $\text{Kt} \equiv E(r_i - \bar{r})^4/\sigma_r^4 - 3$ were calculated to assess the shape of the residual distribution. Table IV lists $|\text{Sk}|/\sigma_{\text{Sk}}$ and $|\text{Kt}|/\sigma_{\text{Kt}}$, where $\sigma_{\text{Sk}}$ is the standard deviation of skewness for a random sample from a normal distribution and $\sigma_{\text{Kt}}$ is the standard deviation of kurtosis for a similar sample. $|\text{Sk}|/\sigma_{\text{Sk}} \leq 2$ and $|\text{Kt}|/\sigma_{\text{Kt}} \leq 2$ are consistent with normal distributions within the 95% significance level, but values exceeding 2 are not consistent within the 95% significance level. Table IV indicates all models exhibit normal residual distributions within $2\sigma$ limits.

V. PARAMETRIC SCALINGS OF PEDESTAL TURBULENCE

In Sec. IV, a search algorithm identified empirical regression models among turbulence quantities and transport-relevant plasma parameters, and the models were screened for proper statistical characteristics. Now, we identify parametric scalings that emerge from model aggregation. Figure 5 shows examples of $x$ distributions for parametric dependencies from model aggregation. $x > 0$ ($x < 0$) indicates the turbulence quantity increases (decreases) at higher parameter values. For example, Figure 5(a) shows $\nabla n_e$ scalings that appear in $24 L_p/\rho_s$ models. The $x$ coefficients cluster around $x \approx 0.3$ despite different constraints and
Finally, nonlinear turbulent transport expressions take the form 

\[ D \sim L_p^2 / \tau_d \]

where \( L_p \) is the poloidal correlation length. Turbulent eddies are anisotropic with \( L_r < L_p \), but we generally expect eddy dimensions to scale proportionately with \( L_r \). Therefore, the random walk model indicates turbulent transport increases at larger \( L_p \) and smaller \( \tau_d \).

The quasi-linear model gives 

\[ D \sim L_p^2 / \tau_d \]

with turbulence quantities, so no weighting scheme is applied in this analysis.

Heuristic turbulent transport models provide general relationships between transport quantities (like diffusivity \( D \) and thermal conductivity \( \chi \)) and turbulence quantities (like \( L_p, k_\theta, \) and \( \tau_d \)). In addition, turbulence theories specify relationships between plasma parameters and turbulent transport models like linear growth rate \( \gamma \), \( D \), and \( \chi \). When combined, transport and turbulence models can specify relationships between turbulence quantities and plasma parameters. For example, the random walk model for isotropic turbulence gives 

\[ D \sim L_p^2 \]

If \( L_p \) regression models indicate \( L_p \propto \nabla n_e(x > 0) \), then the observed scaling is consistent with turbulent transport driven by TEM turbulence because \( \nabla n_e \) is a drive mechanism for TEM turbulence.\(^{29,30}\)

The random walk model for turbulent transport gives 

\[ D \sim L_p^2 / \tau_d \]

where \( L_p \) is the radial correlation length.\(^{31,32}\) Turbulent eddies are anisotropic with \( L_r < L_p \), but we generally expect eddy dimensions to scale proportionately with \( L_r \). Therefore, the random walk model indicates turbulent transport increases at larger \( L_p \) and smaller \( \tau_d \).

The quasi-linear model gives 

\[ D \sim \left( \gamma / k_{\perp,\max} \right) \]

where \( k_{\perp,\max} \) is the characteristic perpendicular wavenumber and \( \gamma \) is the linear growth rate.\(^{33}\) For stationary turbulence (\( k_\perp \sim L_p^{-1} \) and \( \gamma \sim \tau_d^{-1} \)), the quasi-linear model is consistent with the random walk model with 

\[ D \sim L_p^2 / \tau_d \]

Finally, nonlinear turbulent transport expressions take the form 

\[ D \sim L_p^2 / \tau_d \]

where \( \phi \) is the potential perturbation and fluctuation cross-phases are ignored. Turbulent spectra typically follow power laws like 

\[ \phi(k_0) \sim k_0^{-\delta} \]

with \( \delta \approx 2 - 4 \), so nonlinear turbulent transport models are consistent with random walk and quasi-linear models that give 

\[ D \sim 1/k_{\perp,\max} \sim L_p^2 \]

Collectively, heuristic models indicate turbulent transport increases at larger \( L_p \), smaller \( k_\theta \), and smaller \( \tau_d \).

Figure 6 shows parametric scalings for \( L_p, k_\theta, \) and \( \tau_d \) quantities that emerge from model aggregation. Each data point shows the median value and 10th and 90th percentile for the \( x \) distribution like the distributions in Figure 5. The numbers below the plasma parameters indicate the number of models that include the parameter. Notably, plasma parameters that appear for both \( L_p \) and \( k_\theta \) are opposite in sign as expected because \( L \sim k^{-1} \) for broadband turbulence. Also, note that scalings are consistent for different normalizations (e.g., \( L_p, L_p/\rho_i, \) and \( L_p/\rho_s \)). Finally, the analysis excludes \( \nu, \kappa, \delta_0, \rho^s_\nu, \) and \( \rho^s_\kappa \) scalings due to small variation in Table II, as previously mentioned, but the analysis includes other plasma parameters in Table II. Note that not all plasma parameters emerge as good predictor variables. For example, \( \nabla n_e \) and \( \nabla T_i \) are consistent for nonlinear gyrofluid simulations.\(^{31} \)

In Figure 6, the \( \nabla n_e \) and \( \nabla T_i \) scalings for \( L_p \) and \( k_\theta \) are inconsistent with ITG-driven transport. However, \( T_i \) scalings for \( L_p \) and \( k_\theta \) are consistent with ITG-driven transport. Disagreement between \( \nabla n_e \) and \( \nabla T_i \) and \( \nu \) scalings and ITG turbulence theories suggest the observed scalings are inconsistent with ITG-driven transport.

Tapped-electron mode (TEM) turbulence is driven by \( \nabla n_e \) and TEM-driven transport is enhanced at smaller \( \nu, T_i, \) and \( \nabla n_e \) according to nonlinear gyrofluid simulations.\(^{31} \)

In Figure 6, the \( \nabla T_i \), \( \nabla n_e \), and \( \nu^*_i \) scalings for \( L_p \) and \( k_\theta \) are inconsistent with ITG-driven transport. However, \( T_i \) scalings for \( L_p \) and \( k_\theta \) are consistent with ITG-driven transport. Disagreement between \( \nabla n_e \) and \( \nabla T_i \) and \( \nu \) scalings and ITG turbulence theories suggest the observed scalings are inconsistent with ITG-driven transport.

Trapped-electron mode (TEM) turbulence is driven by \( \nabla n_e \) and TEM-driven transport is enhanced at larger \( T_i \) and smaller \( \nu \) and \( T_i \) according to linear and nonlinear gyrokinetic simulations.\(^{29,30}\) However, collisions can destabilize dissipative TEM (DTEM) turbulence.
In Figure 6, the $\nabla n_e$ scalings for $L_p$ and $k_0$ are consistent with TEM-driven transport, as are the $T_e$ and $T_i$ scalings for $k_0$. However, the $1/L_{Te}$ scalings for $\tau_d$ are inconsistent with TEM-driven transport. $\nu_e$ scalings appeared in few re-gression models and exhibited the same sign as $\nu_i^*$ scalings. Collectively, the $\nu$ scalings are consistent with DTEM turbulence, not collisionless TEM turbulence. Agreement between $\nabla n_e$, $T_e$, and $T_i$ scalings and TEM theories suggests the observed scalings are partially consistent with TEM-driven transport.

KBM turbulence is driven by pressure gradients and exhibits a critical $\beta_e$ value for onset according to linear gyro-fluid and gyrokinetic simulations. The $\beta_e$ scalings for $L_p$ and $k_0$ in Figure 6 are consistent with enhanced KBM-driven transport at higher $\beta_e$, but the gradients $\nabla n_e$, $\nabla T_i$, and $1/L_{Te}$ scalings give mixed agreement with regard to KMB-driven transport. The partial agreement indicates KMB turbulence can be considered a candidate mechanism for the observed pedestal turbulence.

Microtearing (MT) mode turbulence is driven by $\nabla T_e$, and MT-driven transport is enhanced at higher $\nu_e$ and $\beta_e$ according to linear and nonlinear gyrokinetics. The $\beta_e$ scalings for $L_p$ and $k_0$ in Figure 6 are consistent with enhanced MT-driven transport at higher $\beta_e$. Again, $\nu_e$ scalings appeared in a few regression models and exhibited the same sign as $\nu_i^*$ scalings. Collectively, the $\nu$ scalings are consistent with enhanced MT-driven transport at higher $\nu$. Finally, the $1/L_{Te}$ scalings for $\tau_d$ are inconsistent with enhanced MT-driven transport at higher $1/L_{Te}$. Like KBM, MT turbulence can be considered a candidate mechanism for the observed turbulence based upon $\nu_e$ and $\beta_e$ scalings despite inconsistency with $1/L_{Te}$ scalings. Note that MT simulations for the NSTX core region indicate BES measurements would be insensitive to MT turbulence, but pedestal simulations give a more complex picture. For instance, pedestal simulations point to turbulence with mixed parity or hybrid mode structures, and MT turbulence in the pedestal exhibits less sensitivity to collisions.

Next, we consider flow shear regulation of turbulence. The $\nabla V_i$ scalings for $L_p$ and $k_0$ are consistent with turbulence reduction by equilibrium flow shear. $E \times B$ flow shear scalings are not available due to challenges with second derivatives of profile quantities, but the scalings do indicate $\tau_d$ decreases at higher $E_i$. The observed $E_i$ scaling is consistent with $E \times B$ flow shear decorrelation of turbulence because higher $E_i$ will increase the $E_i$ gradient throughout the pedestal, as shown in Figure 4. Therefore, the observed $E_i$ scalings for $\tau_d$ are consistent with turbulence decorrelation by $E \times B$ flow shear. Also, the $\nu_i^*$ scalings in Figure 6 point to enhanced turbulent transport at higher $\nu_i^*$, consistent with reduced zonal flow activity at higher $\nu$.

Finally, the observed scalings in Figure 6 show $L_p$ increases at larger pedestal height ($h_{ped}$) and $k_0$ decreases at larger $h_{ped}$ and pedestal width ($A_{ped}$). The scalings are consistent with the link between wider, taller pedestals and larger turbulent structures recently reported for tokamak edge turbulence measurements.

To recap, heuristic transport models, turbulence theories, and the observed scalings in Figure 6 can provide insight into pedestal turbulence mechanisms. The observed scalings show the most consistency with TEM, KMB, and MT-driven transport, but the scalings are least consistent with ITG-driven transport. The $\nabla V_i$ and $E_i$ scalings are consistent with turbulence reduction by flow shear, and the $\nu$ scalings are consistent with zonal flow regulation of turbulence. Also, the scalings are consistent with the link between wider, taller pedestals and larger turbulent structures. The scalings in Figure 6 do not implicate a single turbulence mechanism, but rather suggest that multiple instabilities may be active in the strongly sheared, high pressure gradient pedestal region. The scalings are akin to parameter scans in turbulence simulations, so future comparisons to pedestal turbulence simulations and corresponding scans will help unravel the various instabilities and interactions that impact pedestal structure.

VI. SUMMARY

Confinement projections for ITER and next-step devices benefit from accurate pedestal models, and the ST parameter regime provides an opportunity to enhance confidence in pedestal models. Measurements of low-k turbulence in the steep gradient region of the NSTX H-mode pedestal during ELM-free, MHD quiescent periods reveal broadband turbulence with frequencies up to about 50 kHz. Dimensionless turbulence quantities are consistent with drift-wave turbulence parameters with $L_p/\rho_i \sim 10$, $k_0 \rho_i \sim 0.2$, and $\tau_d/(\alpha/\epsilon_i) \sim 5$. We introduced a model aggregation technique to identify parametric dependencies among turbulence quantities and transport-relevant plasma parameters. Model aggregation is an exploratory technique to identify patterns in multi-dimensional datasets with complex interdependencies. The observed scalings indicate $L_p$ increases and $k_0$ decreases with higher $\nabla n_e$ and $\nu_i^*$ and with smaller $\nabla T_i$. The observed scalings show the most agreement with transport driven by trapped-electron mode, kinetic ballooning mode, and microtearing mode turbulence, and the least agreement with ion temperature gradient turbulence. In addition, the scalings are consistent with turbulence regulation by flow shear and consistent with the observed link between wider, taller pedestals and larger turbulent structures. The observed scalings identified by model aggregation are like parameter scans in turbulence simulations, so model validation and comparison to simulations can be straightforward in future work.

ACKNOWLEDGMENTS

The authors gratefully acknowledge helpful discussions with D. R. Mikkelson and S. J. Zweben. This work was supported by U.S. Department of Energy Grant Nos. DE-FG02-89ER53296, DE-SC0001288, and DE-AC02-09CH11466.