Validation of gyrokinetic simulations of a National Spherical Torus eXperiment H-mode plasma and comparisons with a high-k scattering synthetic diagnostic

Validation of gyrokinetic simulations of an NSTX H-mode plasma and comparisons with a high-k scattering synthetic diagnostic

J. Ruiz Ruiz[1], W. Guttenfelder[2], A. E. White[1], N. T. Howard[1], J. Candy[3], Y. Ren[2], D. R. Smith[4], N. F. Loureiro[1], C. Holland[5], and C. W. Domier[6]

[1]*MIT-Plasma Science and Fusion Center, Cambridge, Massachusetts 02139, USA*
[2]*Princeton Plasma Physics Laboratory, Princeton, New Jersey 08543, USA*
[3]*General Atomics, P.O. Box 85608, San Diego, CA, USA*
[4]*University of Wisconsin-Madison, Madison, Wisconsin 53706, USA*
[5]*Center for Energy Research, University of California, San Diego, La Jolla, California 92093-0417, USA*
[6]*University of California at Davis, Davis, California 95616, USA*

Abstract

A new extensive validation study performed for a modest-beta NSTX NBI-heated H-mode discharge predicts that electron thermal transport can be entirely explained by short-wavelength electron-scale turbulence fluctuations driven by the electron temperature gradient mode (ETG), both in conditions of strong and weak ETG turbulence drive. Quantitative comparisons between high-k fluctuation measurements [D. R. Smith, E. Mazzucato, W. Lee, H. K. Park, C. W. Domier, and N. C. Luhmann Jr., Rev. Sci. Instrum., **79**, 123501 (2008)] and simulations are performed via a novel synthetic high-k diagnostic. Computationally intensive electron-scale simulations featuring an unusually large domain $(L_r, L_\theta) \sim (20, 20)\rho_s$ are shown to be required for accurate deployment of the synthetic diagnostic. Ion thermal transport is shown to be close to neoclassical levels, consistent with stable ion-scale turbulence simulations conducted with GYRO [J. Candy and R. E. Waltz, J. Comput. Phys. **186** (2003) 545]. Electron-scale GYRO simulations are shown to match the thermal power-balance estimates from TRANSP. The frequency spectra characteristics of electron-scale turbulence (spectral peak and width) can be consistently reproduced by the synthetic spectra, but these reveal not to be a critical constraint on the simulation model. The shape of the high-k wavenumber spectrum and the fluctuation level ratio between the strong and weak ETG conditions can also be simultaneously matched by electron-scale simulations within sensitivity scans about the experimental profile values, and result to be great discriminators of the turbulence models analyzed. Together, electron thermal power comparisons and quantitative agreement of electron-scale turbulence spectra give strong evidence supporting electron-scale ETG fluctuations as the main mechanism driving anomalous electron thermal transport in the two outer-core conditions of the modest-beta NSTX H-mode analyzed.

I Introduction

Understanding electron thermal transport is a main priority for spherical tokamaks (STs), where energy losses have been dominated by the electron channel [1, 2]. Good electron energy confinement will be fundamental towards achieving good fusion performance in ITER and future FNSF power plants based on the ST concept [3, 4]. The National Spherical Torus eXperiment (NSTX [5]) and the Mega Ampere Spherical Tokamak (MAST [6]) can both reach unique plasma parameter regimes with high levels of $E \times B$ flow-shear driven by the neutral beams, which added to the high beta and enhanced toroidicity are capable of stabilizing ion-scale instabilities such as the ITG. Consistently, ion heat flux estimates have been observed to be close to the irreducible minimum given by neoclassical theory in NSTX [7–9] and MAST [10]. In

these conditions the energy confinement time τ_E can be assumed to be directly determined by the electron thermal transport. Both machines have reported an empirical energy confinement time scaling inversely proportional with collisionality $B\tau_E \sim 1/\nu$ [7,10], resulting highly favorable for the reduced collisionality regimes of future fusion reactors, and one of the strong reasons to pursue the spherical tokamak path to fusion energy. Elucidating what mechanisms might be responsible for the dominant electron thermal losses, and under what conditions, is essential for understanding and predicting core performance of future fusion devices, both in STs and conventional tokamaks, and is part of the scope of this work.

There is ample evidence in the literature [1,2,11,12] supporting electron temperature gradient (ETG) driven turbulence as one of the major mechanisms responsible for anomalous electron thermal transport in NSTX, depending on the parametric regime of operation. To study ETG turbulence fluctuations, a 280 GHz microwave scattering system (the high-k scattering system) was implemented on NSTX [13] to detect density fluctuations from high-k turbulence. With the goal of performing direct comparisons to high-k turbulence fluctuations, a synthetic diagnostic for high-k scattering was previously implemented [14] using the GTS code [15]. This work provided initial quantitative comparisons of electron-scale frequency spectra with synthetically generated turbulence. However the authors reported on uncertainties and systematic errors affecting the computation of the synthetic spectra. Consequently, quantitative agreement between the experimental and synthetic frequency and wavenumber spectra could not be obtained.

This paper focuses on a previously analyzed NSTX NBI-heated H-mode discharge [16]. Nonlinear simulations carried out with the GYRO code [17] for this discharge suggested that ETG could produce experimentally relevant electron thermal transport levels $P_e^{\text{sim}}/P_e^{exp} \sim 30\%$ in conditions of strongly driven ETG turbulence, underpredicting the experimental values. When ETG turbulence was weakly driven, both ion and electron scale turbulence were predicted to be suppressed by GYRO, producing negligible electron thermal transport: an unsatisfying situation. In an attempt to understand the origins of electron thermal transport, we perform a systematic validation study of nonlinear gyrokinetics spanning ion and electron-scale turbulence by combining transport estimates along with direct comparisons of the frequency and wavenumber spectra from high-k turbulence. Using appropriate validation metrics, the comparisons performed in this work will guide the selection of the simulations that exhibit closest fidelity to the experimental constraints, implying the closest representation of reality.

The rest of this article will proceed as follows: in section II we will describe the background plasma conditions and we will introduce the main micro-instabilities present at ion and electron-scales. In section III we will describe the nonlinear gyrokinetic simulation setup used for comparisons against experimental heat fluxes and turbulence spectra. In section IV we will introduce the validation workflow performed in this study, that combines direct comparisons at three different levels in the validation hierarchy (thermal power, wavenumber spectra and frequency spectra). In section V we will focus on electron thermal power comparisons P_e [MW] originating from turbulence both at ion and electron-scales ($k_\theta \rho_s \leq 1$ and $k_\theta \rho_s \geq 1$ respectively). In section VI we will compare the measured frequency and wavenumber spectra of high-k turbulence with synthetically generated turbulence spectra by the GYRO code. The synthetic comparisons will set strict constraints on the turbulence model and will allow us to discriminate between those that best describe the measured features of interest in the underlying turbulence. In sections VII and VIII we will discuss the main findings emanating from this extensive validation effort of nonlinear gyrokinetics for an NSTX H-mode plasma and possible avenues of future work.

II Description of plasma conditions

NSTX H-mode plasma discharge 141767 was chosen for this study [16]. This discharge featured 2 MW of neutral beam injection (NBI) heating in the time span of interest, producing toroidal rotation levels with characteristic Mach number $M \approx 0.2$, an relatively constant levels of line-integrated electron density ($\langle n_e \rangle \sim 6.10^{15}\text{cm}^{-2}$) and toroidal magnetic field ($B_t \sim 0.5$ T). MHD activity is shown to be quite low (details can be found in Ref. [16]). Figure 1.a) displays the plasma current I_p [MW] corresponding to this plasma discharge, where the flat-top phase is achieved after 0.3 s, settling at $I_p \sim 1.1$ MA. At $t \approx 0.4$ s a

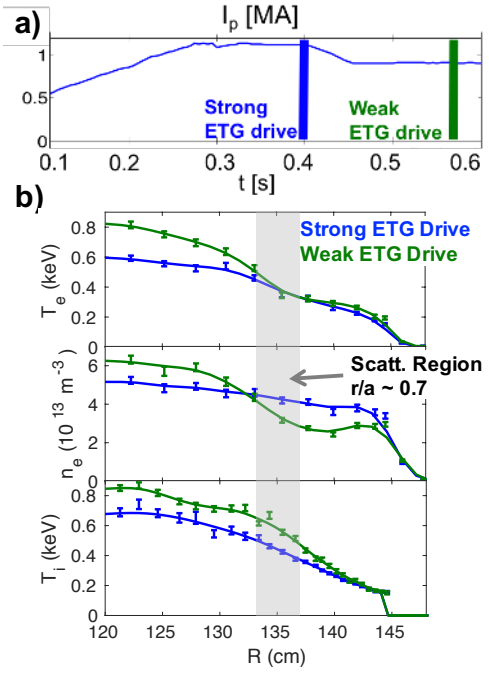

Figure 1: a) Plasma current time trace for NSTX shot 141767. The **Strong ETG drive time** (blue, t=398 ms) takes place during the flat-top phase right before the start of the controlled current ramp-down. The **Weak ETG drive time** (green, t=565ms) takes place after the current ramp-down during a different flat-top period. **b**) Radial profiles of electron temperature T_e, electron density n_e and ion temperature T_i corresponding to the strong and weak ETG drive times. High-k scattering fluctuation data is available at the gray shaded region ($R \approx 135$ cm, $r/a \approx 0.7$).

controlled current ramp-down takes place for about 50 ms, setting a lower plasma current in a new steady phase of $I_p \sim 0.9$ MA. Our validation study will focus on two distinct plasma conditions: before the current ramp-down at $t = 398$ ms will be the **strong ETG drive** condition (blue), corresponding to strong high-k turbulence levels; after the current ramp-down phase at $t = 565$ ms will be the **weak ETG drive** condition (green), corresponding to weak high-k fluctuation levels detected by the high-k scattering system [13].

Figure 1.b) shows the radial profiles of the background electron temperature T_e and density n_e measured using the Multi Point Thomson Scattering system (MPTS [18]), and the ion temperature T_i measured using the Charge Exchange Recombination Spectroscopy sytem (CHERS [19]), corresponding to the strong and weak ETG drive conditions. High-k scattering data from 3 channels is available in the neighborhood of major radius $R \approx 135$ cm ($r/a \approx 0.7$), delimited by the gray shaded region, and will be the main region of interest in this work. The equilibrium electron density gradient displays a large change at the scattering region between the two conditions, changing by a factor of 4 between the two times ($a/L_n \equiv -a/n\nabla n \approx 1$ for the strong ETG case to $a/L_n \approx 4$ for the weak ETG case). As was reported in [16], this large increase in the background electron density gradient was correlated with reduced levels of high-k fluctuation amplitude detected by the high-k scattering system, consistent with the stabilization of ETG by density gradient reported in [12, 20].

The high-k scattering system had 3 operational channels at each time of interest, each sensitive to turbulence wavenumbers $\vec{k}$ in the range $(k_x, k_z) \approx$ (1100-1900, 200-400) m^{-1}, or in normalized values $(k_x\rho_s, k_z\rho_s) \approx$ (7-13, 1.5-3) for the strong ETG condition. Here $\rho_s = c_s/\Omega_{c,i}$ is the ion sound gyro-radius evaluated at the ion sound speed $c_s = \sqrt{T_e/m_i}$ and the ion gyro-frequency $\Omega_{c,i} = eB/m_i$. In this work we denote k_x the wavenumber component along the major radius R, and k_z is component in the vertical direction. The wavenumber component in the toroidal direction k_φ was not used in this work due to the implementation of a 2D synthetic diagnostic (more details in section VII). For the weak ETG condition, $(k_x, k_z) \approx$ (1500-2600, 150-300) m^{-1}, or equivalently $(k_x\rho_s, k_z\rho_s) \approx$ (10-18, 1-2) (the nor-

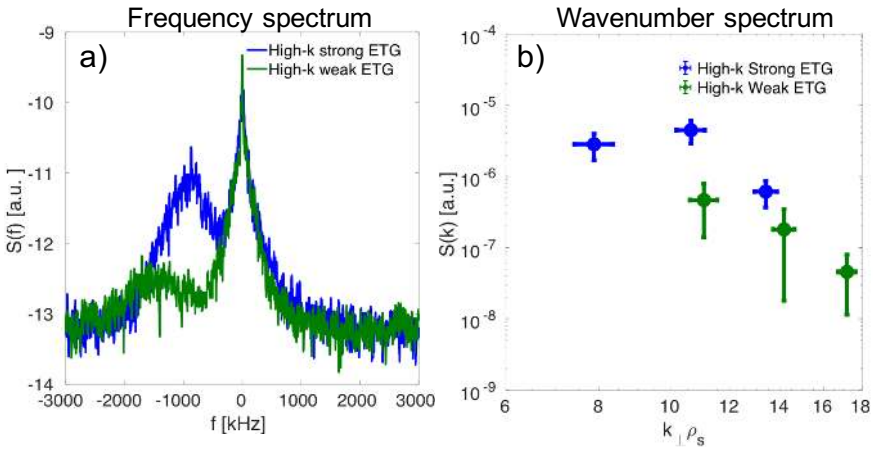

Figure 2: a) Spectral density $S(f)$, or equivalently the frequency power spectrum of electron-scale turbulence fluctuations detected by channel 1 of the high-k scattering system. **b)** Wavenumber spectrum of fluctuations indicative of the total spectral power $S(k)$ from each high-k diagnostic channel. Here $k_\perp$ is defined as the perpendicular component of $\vec{k}_+$ to the background magnetic field. The strong ETG condition is shown in blue while the weak ETG condition is shown in green. Since the high-k diagnostic is not absolutely calibrated, spectra are shown in arbitrary units.

malization by the ion sound gyro-radius $\rho_s \approx 0.7$ cm used TRANSP [21] local values of $T_e \sim 0.4$ keV and magnetic field $B_t \sim 0.5$ T). These wavenumber components are computed independently for each measurement channel of the diagnostic via ray tracing calculations, following the propagation of a single ray modeling the microwave beam propagation in the plasma. Synthetic turbulence spectra will be computed by filtering GYRO-simulated turbulence fluctuations around the measurement wavenumbers by each diagnostic channel, as will be shown in sections IV and VI.

Figure 2 shows the frequency and wavenumber power spectrum of high-k fluctuations detected by the high-k system for the strong and weak ETG conditions. The frequency spectrum in 2.a) is computed from the frequency analyzed density fluctuation signal of channel 1, and used to compute the spectral density $S(f)$. Both spectra exhibit a strong spectral peak at $f = 0$ kHz owing to spurious reflections from the incident microwave beam, resulting in a diagnostic artifact that pollutes the high-k signal. The turbulence spectral power feature is at negative frequency ($f \sim -1$ MHz), and is about an order of magnitude higher for the strong ETG condition with respect to the weak ETG condition (also noticeable from the wavenumber spectrum in 2.b)). The changes in frequency fluctuation power between the strong and weak ETG are due to changes in the turbulence spectrum, but also due to a different wavenumber measurement range for channel 1, as can be seen in b). The wavenumber spectrum is indicative of the total fluctuation power from each wavenumber detected by the high-k system, encoded in the wavenumber spectral density $S(k)$. In b) are plotted channels 1, 2 and 3, channel 1 corresponding to the highest-k value and channel 3 to the lowest-k. The frequency and wavenumber spectra are plotted in arbitrary units due to lack of absolute calibration. More analysis details are given in section IV.

With respect to transport, experimental power balance estimates were computed via TRANSP calculations [21]. A total electron thermal power of $P_e \approx 1.48 \pm 0.33$ MW was obtained for the strong ETG condition, and $P_e \approx 1.02 \pm 0.23$ MW for the weak ETG condition (the $\pm$ sign denotes the 1-σ error bar computed via error propagation analysis). As was discussed in detail in [16], the reduction of ETG fluctuations at the weak ETG time (green in figure 2) was correlated with decreased electron thermal power levels. For reference, ion thermal power $P_i \sim 0.5$-0.6 MW for both conditions, compatible with neoclassical levels predicted by NEO [22] within uncertainty.

Linear stability analysis for the full multiscale spectrum $k_\theta \rho_s \, \epsilon \, [0.1, 100]$ was performed with the gyrokinetic code GYRO in order to understand the main instabilities present for each plasma condition. Experimental profiles taken from TRANSP output were used as input in for the simulations performed at the scat-

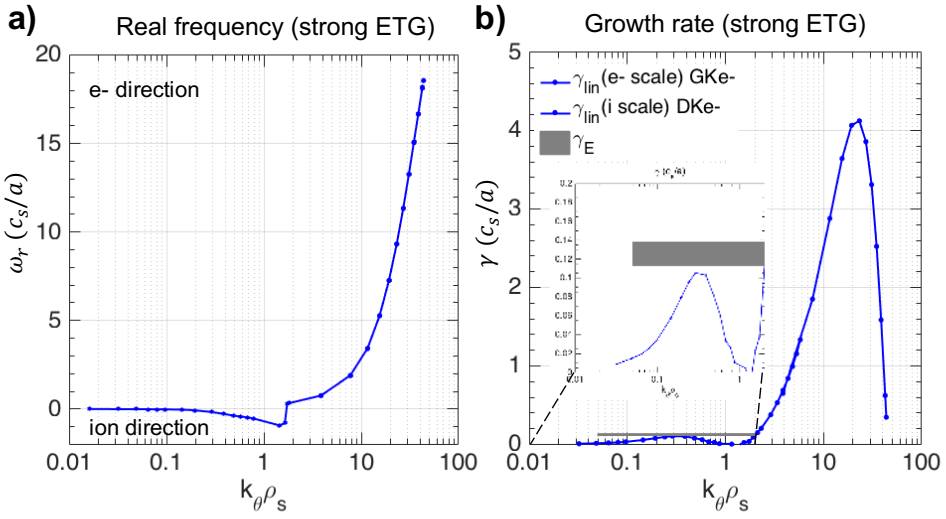

Figure 3: GYRO linear simulation output at strong ETG conditions. **a)**. Real frequency ω_r of instability propagates marginally in the ion direction at low-k ($k_\theta\rho_s \lesssim 1$), and electron directed for $k_\theta\rho_s \gtrsim 1$. **b)** Linear growth rate γ of instability. Zooming on the low-k part of the spectrum shows the $E \times B$ shearing rate γ_E according to the Waltz-Miller definition [23] is greater than the $E \times B$ shearing rate, suggesting the ion-scale instability might be suppressed in this condition (Waltz' rule).

tering location ($r/a \sim 0.7$). Low-k modes were modeled by the drift-kinetic electron approximation, which was shown to agree with the fully gyrokinetic electron approximation up to wavenumber values $k_\theta\rho_s \sim 5$ (not shown), well beyond the expected applicability region and providing confidence in the computed linear growth rates.

The real frequency ω_r and linear growth rate γ for the strong ETG plasma condition are shown in figure 3.a) and b). At ion-scales, the dominant instability is a ballooning-parity mode propagating in the ion diamagnetic drift direction ($\omega_r < 0$) and exhibiting a peak linear growth rate at $k_\theta\rho_s \approx 0.3$. A zoomed snapshot of the instability at low-k (Fig. 3.b) shows that the maximum linear growth rate is smaller than the local $E \times B$ shearing rate for all poloidal wavenumbers $k_\theta\rho_s < 1$, suggesting that the linear instability at ion-scales might be suppressed in this condition [23]. At electron-scales, the dominant linear instability is an electron directed ETG mode ($\omega_r > 0$), exhibiting much higher linear growth rates peaking at $k_\theta\rho_s \approx 20$. This linear gyrokinetic analysis suggests that the ETG mode is a potential candidate to nonlinearly drive experimentally relevant values of electron thermal transport.

Linear stability analysis for the weak ETG condition (Fig. 4) exhibits an electron directed, tearing parity microtearing mode at low-k ($k_\theta\rho_s < 0.1$), transitioning to an electron directed trapped electron mode peaking at $k_\theta\rho_s \approx 0.4$-$0.5$ and enhanced by electromagnetic effects. A range of low-k wavenumbers exhibit linear growth rate values that surpass the local $E \times B$ shearing rate (gray band in figure 4.b)) suggesting the ion-scale instability might survive the stabilization by $E \times B$ shear and drive significant electron heat flux. The ETG peak linear growth rate occurring at $k_\theta\rho_s \approx 100$ suggests the weak ETG condition may drive less electron heat flux than the strong ETG case.

III Nonlinear gyrokinetic simulation setup

Three different types of nonlinear gyrokinetic simulations are discussed in this study: ion-scale gyrokinetic simulation spanning only ion-scale turbulence ($k_\theta\rho_s \lesssim 1$); electron-scale gyrokinetic simulation featuring a conventional domain resolving the ETG mode ($(L_r, L_\theta) = (4.5, 4)\rho_s$); electron-scale simulation with an unusually large simulation domain ($(L_r, L_\theta) = (20, 20.6)\rho_s$). The 'big-box' electron scale simulation was

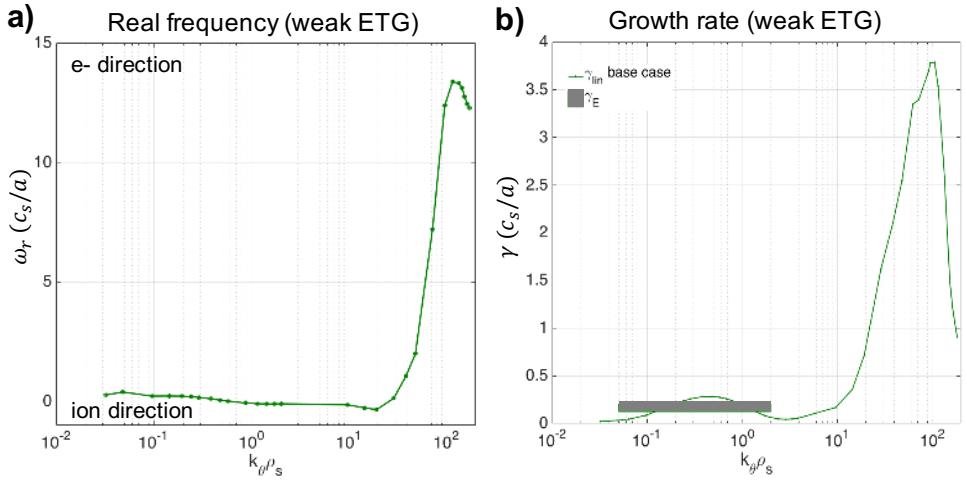

Figure 4: GYRO linear simulation output at weak ETG conditions for the base case (using experimental profile parameters as input). **a)**. Real frequency ω_r of instability propagates in the electron direction at low-k and electron-directed at high-k ($k_\theta\rho_s \gtrsim 10-30$). Note the ion directed mode for $1 \lesssim k_\theta\rho_s \lesssim 30$ transitions to the electron direction for $k_\theta\rho_s \gtrsim 30$. **b)** Linear growth rate γ of instability. Zooming on the low-k part of the spectrum shows the $E \times B$ shearing rate γ_E according to the Waltz-Miller definition [23] is greater than the linear growth rate at all wavenumbers . Note how the linear growth rate at ion-scales is greater than γ_E for the most unstable modes, suggesting ion-scaled turbulence might not be fully suppressed in this condition.

run merely for accurate overlap with the high-k measurement wavenumbers via a synthetic high-k diagnostic, and only accurately resolves the electron-scale turbulence spectrum. The three simulation types will be used to compare turbulence fluxes with TRANSP power balance estimates, but only 'big-box' electron-scale simulation will prove satisfactory for direct high-k turbulence comparisons.

III.a Physics parameters and numerical resolution

Ion-scale simulations model electrons with the drift-kinetic equation, deuterium ions and carbon impurity with the gyrokinetic equation. Standard and 'big-box' electron-scale simulations model all 3 species gyrokinetically. All simulations were performed at the local scattering location $r/a \sim 0.7$, including electron collisions with electron-ion collision frequency $\nu_{ei} = 4\pi n_e Z^2 e^4 \log\Lambda/(2T_e)^{3/2} m_e^{1/2} \approx 1\ \mathrm{c}_s/a$ (via a pitch-angle scattering Lorentz operator). In spite of the relatively low value of the local electron beta

	dr	$L_r[\rho_s]$	n_r	max$(k_r\rho_s)$	$L_\theta[\rho_s]$	d$k_\theta\rho_s$	max$(k_\theta\rho_s)$	n_n	$T[a/c_s]$	$dt[a/c_s]$
ion-scale	$0.3\rho_s$	$70-80$	≈ 200	4	$70-75$	0.08	1.1	14	>200	$2.10^{-3}-10^{-3}$
electron-scale	$2\rho_e$	$6-8$	≈ 200	50.5	$4-5$	1.5	$65-85$	40	30	$10^{-3}-5.10^{-4}$
'big-box' electron-scale	$2.5\rho_e$	$20-25$	≈ 500	$30-40$	$20-22$	0.3	$65-88$	≈ 200	30	$10^{-3}-5.10^{-4}$

Table 1: Numerical resolution parameters typical of ion, electron and 'big-box' electron-scale scale simulation: dr radial resolution (ρ_s, ρ_e are respectively the ion and electron sound gyro-radius using electron temperature T_e), $L_r[p_s]$ radial box size, n_r number of radial modes, max$(k_r\rho_s)$ maximum resolved radial wavenumber, $L_\theta[\rho_s]$ poloidal box size, d$k_\theta\rho_s$ poloidal wavenumber resolution, max$(k_\theta\rho_s)$ maximum resolved poloidal wavenumber, n_n number of toroidal modes, T simulation run time, dt simulation time step. Ion and electron-scale simulations separately resolve ion and electron-scale turbulence. Although 'big-box' electron-scale scale simulation does simulate ion-scale modes ($k_\theta\rho_s < 1$ due to a diagnostic requirement, $cf.$ figure 5), the ion-scale turbulence spectrum is not adequately resolved, neither spatially (which would require need $L_r \gtrsim 70$, d$k_\theta\rho_s \sim 0.08$ as in ion-scale simulation) nor temporally (which would require need $T^{sim} \gtrsim 200$ a/c$_s$ for ion-scale turbulence to reach a fully saturated state). 'Big-box' electron-scale scale simulation only adequately resolves electron-scale turbulence. Accurately simulating both ion and electron-scale turbulence would require multiscale simulations such as shown in [24–28], which is outside of the scope of this work.

$\beta_e \approx 0.003$, fully electromagnetic fluctuations ($\delta\phi, \delta A_{||}, \delta B_{||}$) were kept ($\beta_e$ uses the internal B_{unit} definition in GYRO [17]; normalization with respect to the local magnetic field at the outboard midplane results in $\beta_e \sim$ 3-4%). Simulations include background toroidal flow and flow shear ($M \sim 0.2$, perpendicular flow shear $\gamma_E \sim$ 0.1-0.2, parallel flow shear $\gamma_p \sim 1$), but not a fast ion population due to the small fast ion pressure with respect to the electron pressure ($\lesssim 10\%$). Additional details on the physics parameters input in the nonlinear GYRO simulations can be found in appendix A.

With respect to numerical resolution parameters, linear background profiles were simulated employing non-periodic boundary conditions in the radial direction, implemented in GYRO via buffer damping regions of radial width $\Delta_b \sim 8/1/2\text{-}3\rho_s$ respectively for ion-scale/electron-scale/'big-box' electron-scale simulation. Parallel resolution employed 14 poloidal grid points ($\times 2$ signs of parallel velocity), 12 energies and 12 pitch-angles (6 passing + 6 trapped). This choice of numerical grids was made according to previous convergence and accuracy tests for the GYRO code simulating micro-instabilities in the core of NSTX [1, 11], and was also tested for convergence for the present conditions (appendix B).

III.b Radial and poloidal wavenumber resolution

The radial and poloidal wavenumber resolution of the nonlinear simulations presented is of crucial importance in order to accurately resolve the wavenumbers measured by the high-k system. Ion-scale simulation resolves radial and poloidal wavenumbers characteristic of ion-scale instabilities $k_r\rho_s \in [0.08, 4]$ and $k_\theta\rho_s \in [0.08, 1.1]$. Standard electron-scale simulation resolves only electron-scale turbulence wavenumbers $k_r\rho_s\epsilon$ [1, 30-50] and $k_\theta\rho_s\epsilon$ [1.5, 65-85] depending on the plasma condition. 'Big-box' electron-scale simulation only accurately resolves electron-scale turbulence, but includes modes characteristic of low-k instabilities, $k_r\rho_s \in [0.3, 40]$ and $k_\theta\rho_s \in [0.3, 65\text{-}88]$ depending on the plasma condition. Figure 5 displays the radial and poloidal wavenumber simulation grid from a standard electron-scale simulation (left) and a 'big-box' electron-scale simulation (right), along with typical wavenumbers detected by channels 1, 2 and 3 of the high-k scattering system. The black dots denote the dominant wavenumber detected by each channel, while the ellipses indicate the wavenumber resolution ($1/e$ amplitude of the effective wavenumber filter). Ideally, one would want to simulate several radial and poloidal wavenumbers inside each (k_r, k_θ)-ellipse to accurately replicate the experimental fluctuation measurement. However, due to a coarse wavenumber grid spacing, standard electron-scale simulation can only resolve 1-2 poloidal wavenumbers inside the effective wavenumber measurement range (figure 5.a)). This poor overlap between the diagnostic and numerical resolution requirements results in inaccurate frequency spectra from electron-scale simulation (figure 6). By decreasing the wavenumber grid spacing (dk_r) and (dk_θ), 'big-box' electron-scale simulation can effectively filter $\sim$ 5-6 poloidal wavenumbers inside the measurement range from each channel. This results in a closer representation of the actual high-k measurement. These issues are further discussed in the next section.

III.c Comparisons of standard vs. 'big-box' electron-scale simulation setup

'Big-box' electron-scale simulations reproduce the same physics of ETG-driven turbulence as traditional electron-scale simulations. The essential difference between the two simulation types resides in the numerics, not in the simulated physics. Due to their increased computational demands only a handful of 'big-box' electron-scale simulations were performed in this study: five for the strong ETG condition and two for the weak ETG condition. On the other hand, multiple parametric scans performed for the less-computationally intensive electron-scale simulation provide valuable information about the turbulence threshold, stiffness, etc. and guided the deployment of 'big-box' electron-scale simulations for the specific driving parameters presented. Typical 'big-box' electron-scale simulations presented here have factors of 2-5 more radial grid points n_r and toroidal mode numbers n_n than standard electron-scale simulations, while they typically ran on 10-20 thousand parallel CPU cores taking $\sim$1-2 M CPU hours to completion on leadership high-performance supercomputers such as NERSC's Edison. For additional numerical details, the main numerical simulation parameters differentiating ion-scale, electron-scale and 'big-box' electron-scale sim-

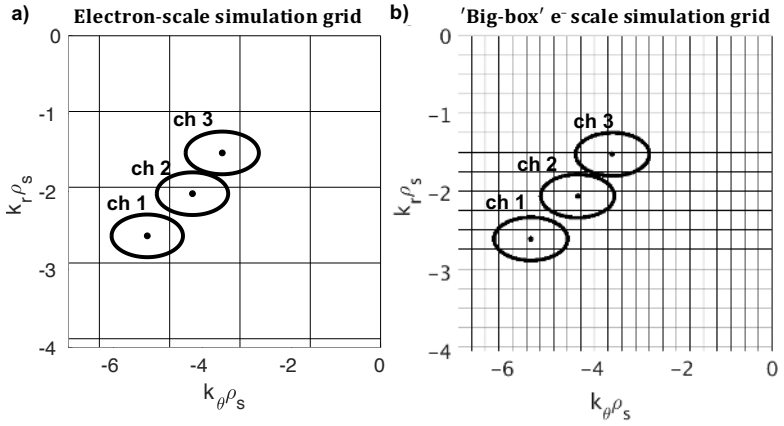

Figure 5: (k_r, k_θ) simulation grid for electron-scale (left) and 'big-box' electron-scale scale simulation (right) along with the measurement range of channel 1, 2 and 3 of the old NSTX high-k scattering system [13]. The ellipses denote the wavenumber resolution of each diagnostic channel (the wavenumber measurement range). Simulation wavenumbers inside the ellipses will effectively contribute to the synthetic signal. **a)** electron-scale simulation does not well resolve the measurement wavenumbers from the high-k diagnostic due to a coarse (k_x, k_θ)-grid. **b)** 'big-box' electron-scale scale simulation is needed to accurately resolve the measurement wavenumbers from the old high-k scattering diagnostic. By resolving more (k_x, k_θ)-modes than a standard electron-scale simulation, 'big-box' electron-scale simulation has a bigger radial and poloidal simulation box, effectively resolving some ion-scale modes (note the k_θ simulation resolution is $\mathrm{d}k_\theta\rho_s^{sim} \sim 0.3$).

ulation can be found in table 1.

In order to justify the use of 'big-box' e- scale simulation over standard e- scale simulation for quantitative spectra comparisons, a frequency power spectrum computed from each simulation type is shown in figure $6.b)$. When compared to the experimental frequency spectrum shown in $a)$, 'big-box' e- scale simulation shows a much better agreement than standard e- scale simulation, as well as an improved agreement in the width of the spectrum (spectral width W_f). In addition, standard e- scale simulation spectra exhibit a 'double-peak' structure not present in experiment. This can be explained via inspection of the simulation wavenumber grid in figure 5. As is shown in figure $5.a)$, the $1/e$ amplitude of the effective filter applied to simulated fluctuations from standard electron-scale simulation has predominantly contributions from two poloidal wavenumbers k_θ. Each poloidal wavenumber has a different characteristic propagation frequency, and is Doppler shifted by a different frequency $f_D = \omega_0 k_\theta r/q$. This results in a clear separation of the propagation frequency in the lab frame, providing a double-peak structure in figure $6.a)$. This frequency separation of the sampled wavenumbers is not observed when using 'big-box' e- scale simulation in $6.b)$ due to the increased number of sampled modes. The increased wavenumber resolution has the effect of 'filling in' the frequency spectrum with contributions from additional nearby modes, and is also responsible for a widening of the spectrum. This shows how the simulation wavenumber resolution can have a profound effect on the frequency spectrum, even for identical physics parameters. We emphasize this argument is purely based on numerical requirements, and is not a reflection of the simulated physics. 'Big-box' e- scale simulations are shown to be required to reproduce the frequency spectrum from experiment.

Note that 'big-box' electron-scale simulation does effectively simulate ion-scale modes due to an increased simulation box with respect to conventional electron-scale simulation (table 1), but does not correctly resolve the full spectrum of ion-scale turbulence (which would require $\mathrm{d}k_\theta\rho_s \sim 0.05$-$0.1$). In addition, 'big-box' e- scale simulation is only run for electron time scales ($T \sim 30\ a/c_s$, when ions have not had time to reach a fully saturated state). Consequently, 'big-box' electron-scale simulation should not be considered as multiscale simulation such as the ones documented in [24–28] by Howard *et al.* and Maeyama *et al.*, but rather as electron-scale simulations having a large simulation domain.

IV Validation workflow

We present here the validation workflow performed in this study. The fidelity of each simulation is quantified via validation metrics for electron thermal power and turbulence spectra comparisons. Experiment/model comparisons will be performed at three different levels in the validation hierarchy [29,30]: electron thermal power comparisons with power balance calculations (TRANSP) will be at the highest level in section V, performed for all simulation types; electron-scale turbulence spectra will only be compared to synthetic spectra from 'big-box' electron-scale simulation, using a synthetic diagnostic for high-k scattering developed for this work. Frequency spectra will be at the lowest level in the hierarchy, while wavenumber spectra comparisons will be at the second hierarchy level, both shown in section VI. The methodology employed for the selection of models that best agree with the hierarchy comparisons presented will also be discussed.

IV.a Synthetic diagnostic description

As is described partly in section II and reference [13], the high-k scattering diagnostic on NSTX was a tangential coherent scattering system sensitive to fluctuations with high-k_x and intermediate k_z. Each diagnostic channel is sensitive to an array of wavenumbers that is centered around the dominant measurement wavenumber, according to the wavenumber resolution (ellipses on figure 5). Due to the enhanced toroidicity and magnetic field curvature of NSTX, good toroidal localization is achieved [31, 32]. This supports the implementation of a 2D synthetic diagnostic, which is applied a posteriori to simulation output. Carefully designed 'big-box' electron-scale gyrokinetic simulation accurately resolving each measurement wavenumber are used for the deployment of the synthetic diagnostic.

In order to compute the synthetic time signal of fluctuations, we filter the GYRO-simulated fluctuations around the measurement wavenumber corresponding to each diagnostic channel. This procedure can be performed equivalently in real space or in Fourier space (ie. wavenumber space). Two synthetic diagnostics have been implemented for the purpose of this work, in real space and in k-space, exhibiting quantitative agreement and providing confidence in the accuracy of the synthetic turbulence spectra. In what follows we will only present the real space implementation, leaving the wavenumber implementation and further details to a future publication. The computation of synthetic frequency and wavenumber spectra starts with the synthetic time signal of fluctuations $\delta n_{\mathrm{u}}(\vec{k}_+, t)$, given by

$$\delta n_{\mathrm{u}}(\vec{k}_+, t) = \int_{V_s} d^3\vec{r'}\, \delta n^{\mathrm{GYRO}}(\vec{r'}, t) U(\vec{r'}) e^{-i\vec{k}_+ \cdot \vec{r'}} \tag{1}$$

where the integration takes place over the scattering volume V_s. $U(\vec{r})$ is the effective real space filter of fluctuations delimiting the scattering volume, typically taken to be Gaussian in the R and Z directions owing to the Gaussian profile of the incident microwave beam. $\vec{k}_+$ corresponds to the matched wave-vector between the incident and scattered wave-vectors $\vec{k}_+ = \vec{k}_s - \vec{k}_i$, providing the dominant contribution to the scattered signal (the second matching wave-vector $\vec{k}_- = \vec{k}_s + \vec{k}_i$ gives a negligible contribution to the synthetic signal). We call $\vec{k}_+$ the measurement wavenumber of the diagnostic, assumed constant within the scattering volume in this work. δn^{GYRO} is the GYRO computed density fluctuation field in real space (a real quantity itself). In our 2D implementation the integration takes place at a constant toroidal slice (comments on the 2D $vs.$ 3D implementation are the object of a near-future publication). Real-space integration and filtering of δn^{GYRO} using equation 1 provides a time signal of the filtered fluctuations $\delta n_{\mathrm{u}}(\vec{k}_+, t)$.

IV.b Computation of synthetic turbulence spectra

The time signal of fluctuations $\delta n_{\mathrm{u}}(\vec{k}_+, t)$ can be frequency analyzed to provide the synthetic frequency power spectrum for each diagnostic channel $S(f)$ (spectral density). Integration of the spectral density

$S(f)$ along a prescribed frequency band for the different diagnostic channels will provide the wavenumber spectrum of fluctuations $S(k)$.

An example of high-k frequency power spectrum $S(f)$ from channel 1 of the high-k scattering diagnostic is shown in figure 6.a). A central high power peak is obtained at $f \approx 0$, owing to spurious reflections from the incident microwave beam, resulting in a diagnostic artifact polluting the high-k signal. At a lower power level a negative frequency feature peaking at $f \approx -1$ MHz corresponds to detected high-k turbulence fluctuations. Integrating the spectral density $S(f)$ over the full frequency spectrum would yield a total power dominated by the central spectral peak, hiding crucial information about the turbulence spectral power. One can obtain a measure of the spectral power of the turbulent fluctuations by integrating over a prescribed frequency band, delimited by the black vertical dashed lines. The choice of this frequency band is made to minimize the effect of the background $f = 0$ kHz spectral noise peak while still preserving the bulk turbulence spectral power. A gaussian profile can be fit to the fluctuation frequency spectrum (red curve) to provide estimates for the total spectral power (constituting the wavenumber spectrum $S(k)$), spectral frequency peak $< f >$ and spectral width W_f.

One important aspect to consider is that the experimental signal level is not absolutely calibrated, and can only provide relative fluctuation power. For this reason, the experimental frequency and wavenumber spectra will be scaled by a constant value in order to perform direct quantitative comparisons with synthetic spectra. The scaling constant applied is then a free parameter whose value will be set on each condition, as will be specified in the next section.

IV.c Validation metrics

In order to identify simulations that best satisfy the experimental constraints, we employ simulation/experiment comparisons on three validation hierarchy levels [29, 30]. Electron thermal power comparisons are at the highest level, performed for all simulation types in section V. Local profile values will be scanned around the most significant input drive terms: ∇T and ∇n are varied within 1-σ uncertainty (computed from uncertainties in the background electron temperature and density profiles using a Monte Carlo approach), while q is varied by 10% and $\hat{s}$ is varied by 20% (constituting reasonable errors in the magnetic equilibrium). The specific drive term values are provided in next section. These local transport comparisons will select the 'big-box' electron-scale simulations that best match the electron thermal power (P_e [MW]) within experimental uncertainty. For the strong ETG case, five 'big-box' electron-scale simulations producing experimentally relevant electron thermal transport P_e will be used for further detailed comparisons of high-k spectra. For the weak ETG condition, two 'big-box' electron-scale simulations will be performed, but only one of them will be able to match the electron thermal power. Frequency spectra comparisons will be at the lowest level and shown in section VI, providing detailed information about the spectral peak $< f >$ and spectral width W_f. However, these will not prove useful for the purpose of model selection. Wavenumber spectra comparisons will be at the second level in the hierarchy, and will yield comparisons of the fluctuation level ratio and the k-spectrum shape. Wavenumber spectra comparisons will prove extremely useful for discriminating against turbulence models that best match the experimental constraints, and will be performed for the five 'big-box' electron-scale simulations at the strong ETG and one flux-matching simulation at weak ETG.

We choose the following observables $\mathbf{X}$ for multi-level hierarchy comparisons of the 'big-box' electron-scale simulation models: electron thermal power $\mathbf{P_e}$, fluctuation level ratio $\mathbf{ratio}$ and wavenumber spectra shape $\mathbf{shape}$, defined as

Electron thermal power	$\mathbf{X} = \mathbf{P_e}$	obtained from TRANSP and nonlinear GYRO simulation	
Fluctuation level ratio	$\mathbf{X} = \mathbf{ratio}$	$= \log_{\mathbf{10}} \left(\frac{\langle \mathbf{S(k_i)} \rangle^{\mathbf{strong\ ETG}}}{\langle \mathbf{S(k_i)} \rangle^{\mathbf{weak\ ETG}}} \right)$	ratio of mean spectral densities
k**-spectrum shape**	$\mathbf{X_i} = \mathbf{shape(i)} = \mathbf{S(k_i)}$		spectral density from each diagnostic channel **i** = {1, 2, 3}

(2)

where $\langle \mathbf{S(k_i)} \rangle = \sum_i S(k_i)/3$ denotes the average fluctuation power over the different channels **i** computed using the spectral density $S(k)$ (respectively for the strong and weak ETG conditions in order to compute the ratio). The absence of absolute calibration of the high-k scattering diagnostic power motivates the use of the fluctuation level ratio between the strong and weak ETG conditions as a meaningful metric. The quantity $\frac{\langle \mathbf{S(k_i)} \rangle^{\mathbf{strong\ ETG}}}{\langle \mathbf{S(k_i)} \rangle^{\mathbf{weak\ ETG}}}$ was observed to vary by 2 orders of magnitude and motivated the choice of the function $\log_{\mathbf{10}}$ as a metric measure. To characterize the shape of the wavenumber spectrum via $\mathbf{shape}$, the value of the density fluctuation wavenumber power spectrum $\mathbf{S(k_i)}$ corresponding to each diagnostic channel is employed ($\mathbf{i} = \{1, 2, 3\}$). The observable **shape** is thus a vector of three components $\mathbf{X_i}$.

Each of the three observables ($\mathbf{P_e}$, $\mathbf{ratio}$ and $\mathbf{shape}$) use predictions from simulation $\mathbf{X}^{\text{sim}}$, and are directly compared to experiment $\mathbf{X}^{\text{exp}}$ via a validation metric of distance $\mathbf{d_X}$ (a scalar quantity) that quantifies the goodness of agreement (following methodologies described in [33,34]). In order to have a bounded measure of agreement between 0 and 1, the quantity $\mathbf{R_X}$ is also computed from $\mathbf{d_X}$. Both quantities are defined in past validation works [33, 34] as

$$\begin{aligned} \mathbf{d_X} &= \sqrt{\frac{(\mathbf{X^{exp}} - \mathbf{X^{sim}})^2}{\sigma(\mathbf{X^{exp}})^2 + \sigma(\mathbf{X^{sim}})^2}} \\ \mathbf{R_X} &= \frac{1 + \tanh((\mathbf{d_X} - d_0)/\lambda)}{2} \end{aligned} \tag{3}$$

where we employed characteristic values of $d_0 = 1.5, \lambda = 0.5$ [33, 34]. Since $\mathbf{shape}$ is a 3-dimensional vector, an individual value of $\mathbf{d_{shape(i)}}$ is computed for each component **i**, according to equation 3. Due to the absence of absolute calibration, a constant of proportionality C is applied to the synthetic spectra. Here C is treated as a free parameter, and the final value of $\mathbf{d_{shape}}$ is the minimum value over C of the square root of the sum of the squares from each component, ie. $\mathbf{d_{shape}} = \min_C \sqrt{\sum_i \mathbf{d^2_{shape(i)}}}$. This is equivalent to the least-squares distance between experiment and simulation, and agrees with the intuition behind the shape of the wavenumber spectrum $S(k)$: a low $\mathbf{d_{shape}}$ value will correspond to a well matched wavenumber spectrum shape, while large $\mathbf{d_{shape}}$ will correspond to a poorly matched shape.

Using the bounded metric $\mathbf{R_X}$ from each of the three observables, a composite metric $\mathbf{M_1}$ can be constructed to quantify the overall fidelity of the experiment/model comparison. A value of 0 in $\mathbf{M_1}$ and $\mathbf{R_X}$ indicates perfect agreement, while 1 is indicative of complete disagreement. The $\mathbf{M_1}$ composite metric will prove useful for discerning and discriminating the different models that best agree with the experimental constraints, indicating the 'best' model representative of reality.

V Local transport comparisons via gyrokinetic simulation

We present in this section electron thermal power predictions (P_e [MW]) from nonlinear gyrokinetic simulations using the GYRO code, separately resolving ion and electron-scale turbulence. Parametric scans were performed for the main turbulence drive mechanisms ($a/L_{Te}, a/L_{ne}, q, \hat{s}$), $a/L_{Te} \equiv -a\nabla T_e/T_e$ and $a/L_{ne} \equiv -a\nabla n_e/n_e$ are the background electron temperature and density gradients respectively, q is the magnetic safety factor and $\hat{s}$ is the magnetic shear (we will note $a/L_{Te}, a/L_{ne}$ interchangeably by $\nabla T, \nabla n$ in the rest of this paper).

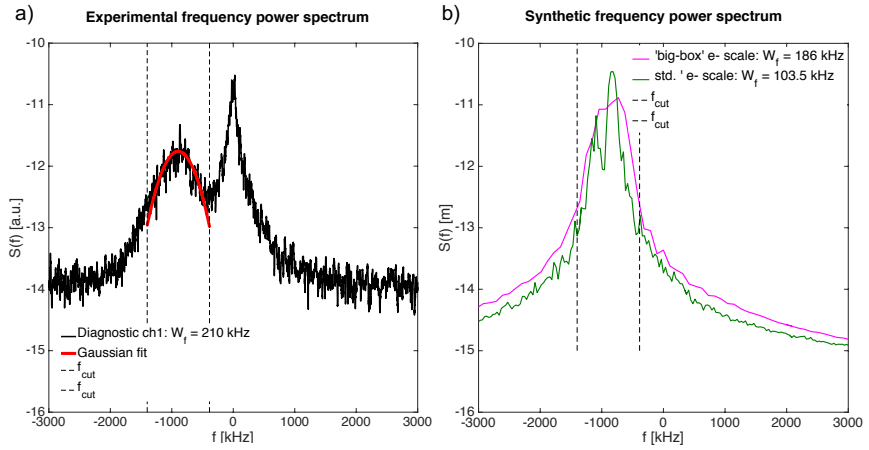

Figure 6: Qualitative comparison of the frequency power spectrum from the high-k diagnostic in **a)** and from simulation in **b)**, both corresponding to channel 1. A gaussian profile is fit within the prescribed frequency bands (vertical dashed lines) to recover the fluctuation power level, the spectral frequency peak of fluctuations $< f >$ and the spectral width W_f (red curve in a)). In **b)** are compared the synthetic spectra from a 'big-box' $vs.$ standard e- scale simulation. The shape of the frequency spectrum is better reproduced by the 'big-box' e- scale simulation, as well as the calculated spectral width W_f, motivating the use of 'big-box' e- scale simulation for establishing quantitative comparisons with experiment. The 'double-peak' structure using standard e- scale simulation can be explained by the Doppler-shifted frequency of the two individual, simulated poloidal wavenumbers k_θ within the k-measurement range ($cf.$ figure 5.a)).

V.a Strong ETG drive conditions

It was discussed in the introduction that electron-scale gyrokinetic simulation using experimental profile values as input can only provide $\approx 30\%$ of the experimental electron thermal power estimate from TRANSP, where $P_e^{exp} \approx 1.48 \pm 0.33$ MW. Ion-scale turbulence simulated by GYRO is shown to be suppressed by the large value of $E \times B$ shear, consistent with the linear simulation results of figure 3. This clear under-prediction of P_e^{exp} leads us to perform parametric scans for the main turbulence drive mechanisms, both at electron and ion-scales: the background electron temperature gradient $a/L_{Te} \approx 3.36$ is the ETG turbulence drive mechanism, and a 1-σ uncertainty corresponds to 25% of the background value. The electron density gradient a/L_{ne} is shown to be a stabilizing mechanism ([16]), and a 1-σ uncertainty corresponds to 50% of the background value (note a/L_{ne} is small for this condition, ≈ 1). The uncertainties in the profile gradients $\sigma(\nabla T)$, $\sigma(\nabla n)$ were computed from uncertainties in the background electron temperature and density profiles followed by a Monte Carlo analysis approach.

Figures 7.a) and b) show the electron thermal power P_e predictions by ion-scale (red), electron-scale (blue) and 'big-box' electron-scale simulation (magenta) when the experimental value of normalized equilibrium electron temperature gradient ∇T is scanned within experimental uncertainty. The experimental value, as computed by TRANSP, is shown in black. Figure 7.a) shows simulations run using the experimental value of the background electron density gradient ∇n as input. Figure 7.b) shows simulations that were run using a reduced value of the background density gradient within 1-σ experimental uncertainty in the input, in order to maximize the turbulence drive.

Figures 7.a) and b) show that ion-scale simulation predicts negligible electron thermal power for all parametric scan values in ∇T and ∇n (red dots), consistent with turbulence suppression by $E \times B$ shear. However, electron-scale simulation can match P_e^{exp} in figure 7.b), using scaled values of ∇n and ∇T (blue dots). 'Big-box' electron-scale simulation provides similar P_e predictions as electron-scale simulation, essentially matching P_e^{exp} when temperature and density gradients are scaled simultaneously (black dot in 7.b)). Due to possible errors in the magnetic equilibrium reconstruction, additional scans with reduced safety factor q by 10% and increased magnetic shear $\hat{s}$ by 20% were performed with 'big-box' electron-scale simulation: the purple square in 7.a) corresponds to a flux-matching simulation with scaled (∇T)

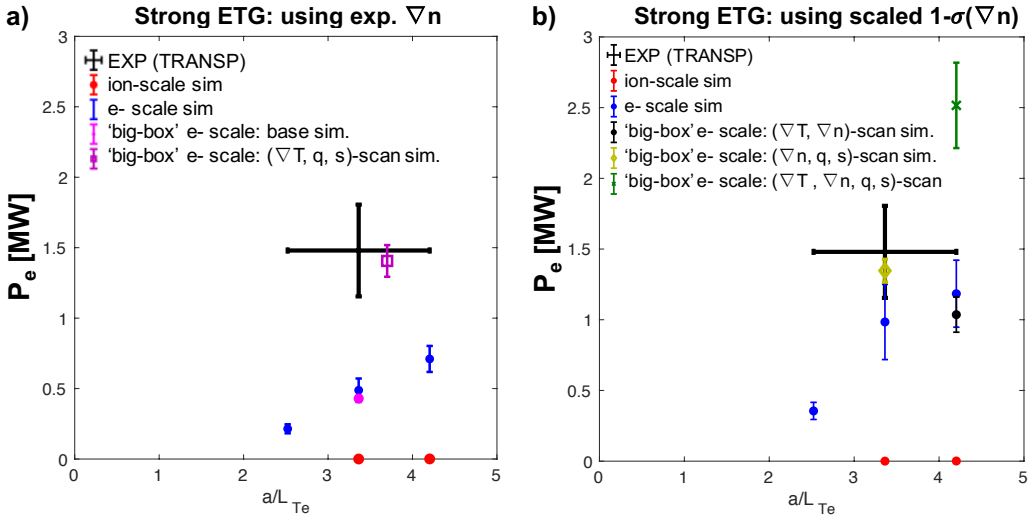

Figure 7: a) Total thermal transport budget from experiment (TRANSP, black) along with GYRO simulations that use experimental value of ∇n as input, corresponding to the strong ETG drive condition. The magenta dot corresponds to 'big-box' electron-scale simulation run for base experimental parameters. The purple square corresponds to the 'big-box' electron-scale simulation with scaled (∇T) within 1-σ uncertainty in addition to -10% q and $+20\%\hat{s}$. **b)** GYRO simulations have scaled ∇n by 1-σ to maximize turbulence drive $(-\sigma(\nabla n))$. Black dot corresponds to 'big-box' electron-scale scale simulation scanning $(\nabla T, \nabla n)$ within uncertainty. Green diamond scans (∇n) within uncertainty, -10% q and $+20\%\hat{s}$. Dark green cross scans $(\nabla T, \nabla n)$ within uncertainty, -10% q and $+20\%\hat{s}$. These simulations show that ion-scale turbulence is stabilized by $E \times B$ shear and electron scale turbulence can explain the experimental electron thermal transport levels within small variations in the input drive terms.

within 1-σ uncertainty ($a/L_{Te} = 3.7$) in addition to -10% q and $+20\%$ $\hat{s}$; the green diamond in 7.b) is also flux-matching, and uses scaled (∇n) within uncertainty (-1-σ), -10% q and $+20\%$ $\hat{s}$; the dark green cross in 7.b) overpredicts P_e, and uses scaled $(\nabla T, \nabla n)$ within uncertainty, -10% q and $+20\%$ $\hat{s}$. Among the five 'big-box' electron-scale simulations in figure 7, two of them can reproduce the experimental P_e value within error-bars, providing strong evidence pointing towards electron-scale turbulence as the main mechanism responsible for anomalous electron thermal transport in this condition and consistent with previous studies in NSTX [1] and MAST [35]. Ion thermal power P_i was shown to be close to neoclassical levels, as predicted by NEO neoclassical calculations [22] (consistent with suppressed ion-scale turbulence), and negligible ion thermal power was predicted by all electron-scale and 'big-box' electron-scale scale simulations. In the next section, synthetic spectra comparisons will be performed for the five 'big-box' electron-scale scale simulations of figures 7.a) and b).

V.b Weak ETG drive conditions

As was discussed in the introduction, the weak ETG condition is especially interesting from a validation perspective, since neither ion nor electron-scale gyrokinetic simulations could produce experimentally relevant values of electron thermal power when the simulation is run using the experimental profile values as input. As mentioned, the TRANSP power balance estimate $P_e^{exp} \approx 1.02 \pm 0.23$ MW.

Similar to the approach followed for the strong ETG condition, parametric scans varying electron temperature gradient input a/L_{Te} were performed within experimental uncertainty, using the experimental density gradient value in figure 8.a), and a fixed scaled value of density gradient of 1-σ uncertainty in figure 8.b). Ion-scale and electron-scale simulations are respectively red and blue, while 'big-box' electron-scale simulations have different colors according to the parameters scanned. The 1-σ uncertainties for a/L_{Te} and a/L_{ne} are respectively 20% and 30% of the background value, computed using uncertainties in the background electron temperature and density profiles followed by a Monte Carlo analysis approach.

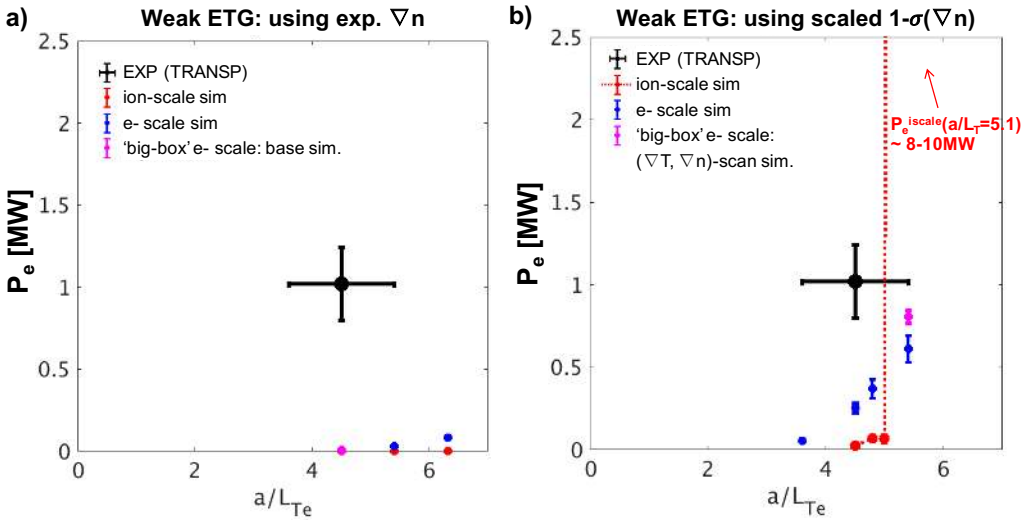

Figure 8: a) Total thermal transport budget from experiment (TRANSP, black) along with GYRO simulations that use the experimental value of ∇n as input, corresponding to the weak ETG drive condition. **b)** GYRO simulations with ∇n scaled by 1-σ to maximize turbulence drive ($-\sigma(\nabla n)$). Note the significant increase in stiffness predicted from electron-scale turbulence (blue) and especially ion-scale turbulence (red) when ∇n is scaled by 1-σ.

Figure 8.a) shows that negligible electron thermal transport is obtained when the background density gradient is set to the experimentally measured value, neither from ion nor from electron-scale turbulence. On the other hand, when a temperature gradient scan is performed in addition to a 1-σ scaled value of the electron density gradient as in figure 8.b), a substantial increment in the predicted P_e can be observed. For values of $a/L_{Te} < 5$, the turbulence state is shown to be dominated by large zonal flow amplitudes suppressing the ion-scale fluctuations and transport. For $a/L_{Te} > 5$ ion-scale simulation predicts values of P_e up to factors of 10 larger than the experimental levels. The transition from a zonal flow dominated state for $a/L_{Te} < 5$ to a turbulent state dominated by large amplitude turbulent fluctuations happens abruptly near $a/L_{Te} \approx 5$, displaying large stiffness. In fact, ion-scale simulations run for values of $a/L_{Te} = 4.5, 4.8, 5$ display zonal flow dominated time phases producing negligible transport, followed by sudden turbulence bursts where lone large scale eddies produce instanteneous high levels of transport. These large-scale eddies are eventually damped by the strong zonal flows present, producing negligible time-averaged electron thermal power as shown in 8.b). This behavior is reminiscent of a Dimits shift-like turbulence regime close to the nonlinear threshold and mediated by the large amplitude zonal flows [36], and more recently observed in conditions of subcritical ITG turbulence enabled by $E \times B$ flow shear [37]. However, for $a/L_{Te} > 5$, the turbulence drive from the background gradients is shown to be too strong for zonal flows to damp a highly unstable ion-scale mode (possibly a hybrid trapped electron mode/kinetic ballooning mode TEM/KBM). This highly unstable mode is strongly sensitive to $\delta B_{||}$ fluctuations, exhibiting complete turbulence stabilization when $\delta B_{||}$ fluctuations were not included in the simulation. Linear β_e scans not presented here showed the KBM β_e threshold was close to the experimental β_e value (within $\sim 20\%$, not shown). Previous work [1] already reported on a hybrid TEM/KBM mode being capable of driving substantial experimental thermal transport which could possibly be linked to this highly unstable ion-scale turbulence regime.

Figure 9 shows the ion and electron thermal power from ion-scale simulation. The predicted ion thermal power is negligible for $a/L_{Te} < 5$ (similarly to the electron thermal power), but overpredicts the experimental TRANSP value by factors of $\sim$ 2-3 (magenta curve) for $a/L_{Te} > 5$. Since the TRANSP P_i is close to the neoclassical NEO prediction, we conclude that ion-scale turbulence should produce negligible transport (P_e and P_i), and the TEM/KBM-stable regime ($a/L_{Te} < 5$) is the most experimentally relevant at the weak ETG condition.

Electron-scale turbulence is also shown to be close to the nonlinear threshold in this plasma regime. As is shown by the blue dots in figure 8.b), a linear dependence of P_e with a/L_{Te} shows a clearly active ETG

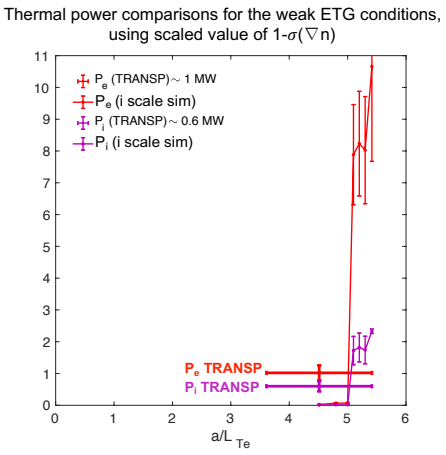

Figure 9: Total thermal power on the electron and ion channel (P_e and P_i respectively) as a function of driving mechanism a/L_{Te} for the weak ETG drive condition. The density gradient scale length value a/L_{ne} in the simulations presented in this figure has been scaled by 1-σ of the experimental uncertainty. Note the increase in predicted P_e and P_i for $a/L_{Te} \sim 5$, displaying large stiffness from a highly unstable TEM branch.

mode producing experimentally relevant P_e values that are close to matching the experimental TRANSP levels for the highest temperature gradient scanned and exhibiting finite stiffness. A 'big-box' electron-scale simulation was run for this highest drive condition (magenta dot), matching P_e^{exp} and slightly overpredicting the electron-scale simulation value for the same scaled a/L_{Te} and a/L_{ne} values. This slight mismatch between electron and 'big-box' electron-scale simulation P_e predictions could be due to the inclusion of highly unstable ion-scale modes down to $k_\theta \rho_s \sim 0.3$ (as shown by ion-scale simulation in red). However those are not fully developed and have not reached a fully saturated state due to the lack of resolution (both temporal and spatial) in 'big-box' electron-scale simulation resolving ion-scale modes. 'Big-box' electron-scale simulation run for the base case is shown in figure 8.a), producing negligible P_e.

These results suggest that ETG rather than ion-scale TEM/KBM is responsible for electron thermal transport. This is consistent with 'big-box' electron-scale simulations being able to match P_e^{exp} within experimental uncertainty in $(\nabla T, \nabla n)$, while the neoclassical ion-thermal transport constraint suggests that the TEM/KBM stable regime is the most experimentally relevant. To test whether electron-scale turbulent fluctuations predicted by GYRO are consistent with the turbulence spectra measurements from the high-k diagnostic, we set out to establish quantitative turbulence comparisons in the next section.

V.c Assessment of cross-scale coupling

Although multiscale simulation is not performed in this work, based on the current single scale simulations we provide here a discussion to assess the potential importance of cross-scale coupling interactions (between ion and electron scales) in the conditions analyzed.

In the strong ETG drive case, ion-scale turbulence is strongly stabilized by the background $E \times B$ flow shear and electron-scale turbulence alone can explain the experimental electron transport levels. Cross-scale coupling is not expected to play an important role. In the weak ETG drive case, ion-scale simulation predicts the experimental conditions to be in a zonal flow dominated regime and very close to the nonlinear threshold, exhibiting a large stiffness value (Fig. 9). Interestingly, electron-scale turbulence is also shown to lie very close to the nonlinear threshold. There exists a regime close to $a/L_{T_e} = 5$ in which electron-scale turbulence is highly driven while ion-scale turbulence is very close to the threshold (Figure 8.b)). These conditions are very similar to those reported by Howard *et al.* in [26, 27] as part of the multiscale simulation efforts for some C-Mod L-mode discharges. These showed that the electron thermal power level can be enhanced by cross-scale coupling interactions between ion-scales and electron-scales when ion-scale

turbulence is close to threshold while electron-scale turbulence is highly driven. This would suggest that cross-scale coupling between ion-scale and electron-scale fluctuations might be important in the weak ETG condition.

In similar lines, previous work by G. Staebler *et al.* [38] indicated the value of $(\gamma/k_\theta)_{ion}/(\gamma/k_\theta)_{ETG}$ is a proxy for testing the importance of multiscale interactions, where γ is the peak linear growth rate and k_θ is the corresponding wavenumber, respectively for the ion-scale and electron-scale spectrum (denoted by the subscripts ion and ETG). Staebler $et\ al.$ found that cross-scale interactions tend to be important for values of $(\gamma/k_\theta)_{ion}/(\gamma/k_\theta)_{ETG} < 1$. In the present conditions we find $(\gamma/k_\theta)_{ion}/(\gamma/k_\theta)_{ETG} \sim 1.9$ for the strong ETG case and $(\gamma/k_\theta)_{ion}/(\gamma/k_\theta)_{ETG} \sim 18$ for the weak ETG (figures 3 and 4). This would suggest that cross-scale interactions are not important, even less for the weak ETG condition. It may also be that this rule of thumb, which was implemented based on C-Mod L-mode plasmas, is not applicable for the present conditions and would need further revision for spherical tokamak conditions. The question remains open and only multiscale simulation would confirm whether cross-scale interactions are relevant in these conditions.

VI Synthetic diagnostic comparisons of high-k fluctuation spectra

Direct comparisons between experiment and simulation of high-k frequency and wavenumber spectra are presented in this section. Only 'big-box' electron-scale simulations will be used to deploy the synthetic diagnostic for high-k scattering. For the strong ETG condition, we will be use the five 'big-box' electron-scale simulations shown in figures 7.a) and b): 'big-box' base sim (magenta in a)), 'big-box' $(\nabla T, q, \hat{s})$-sim (purple square in a)), 'big-box' $(\nabla T, \nabla n)$-sim (black dot in b)), 'big-box' $(\nabla n, q, \hat{s})$-sim (green diamond in b)), 'big-box' $(\nabla T, \nabla n, q, \hat{s}$-sim (dark green cross in b)). For the weak ETG condition, we will only use two 'big-box' electron-scale simulations shown in figures 8.a) and b): 'big-box' base sim (magenta in a)) and 'big-box' $(\nabla T, \nabla n)$-sim (magenta in b)).

VI.a Frequency spectra comparisons

Following the procedure outlined in section IV one can extract the frequency spectral peak $< f >$ and spectral width W_f and use them as validation metrics. Figures 10.a), b) and c) show frequency spectra comparisons for all channels of the strong and weak ETG case (blue and green respectively), using simulations that matched the electron thermal power. Note that the synthetic frequency spectra was computed implementing the full Doppler shift value. Experimental frequency spectra were computed within time windows of 546μs (≈ 120 a/c$_s$), and containing 8,192 time points. Due to the large simulation requirements, synthetic frequency spectra were computed in time windows of 15 a/c$_s$ for the strong ETG condition and 22 a/c$_s$ for the weak ETG condition, typically containing 1,500-2,200 time points. It was found that frequency spectra characteristics had little variation with time windows larger than ≈ 15 a/c$_s$ for these conditions. In order to reduce simulation noise, the synthetic frequency spectra was ensemble averaged using five time windows for the strong ETG case (five time windows of ~ 3 a/c$_s$), and seven time windows for the weak ETG case (seven time windows of ~ 3 a/c$_s$).

Good agreement in the frequency spectra is obtained both for the strong and weak ETG conditions and for all operating channels, as is shown in figure 10 and tables 2 and 3. Particularly good agreement is found in the spectral peak of fluctuations $< f >$ which lie in the range $< f > \sim -1$-1.5 MHz, typically within 5% of the experimental spectral peak and much higher than the values obtained when no Doppler shift is applied (figure 11).

The spectral width W_f yields close agreement to experiment for the strong ETG condition, within 20% of the experimental range (table 2), a value marginally within the experimental uncertainty value. However W_f tends to be slightly underpredicted for the weak ETG condition (up to 40-50%, table 3). This could

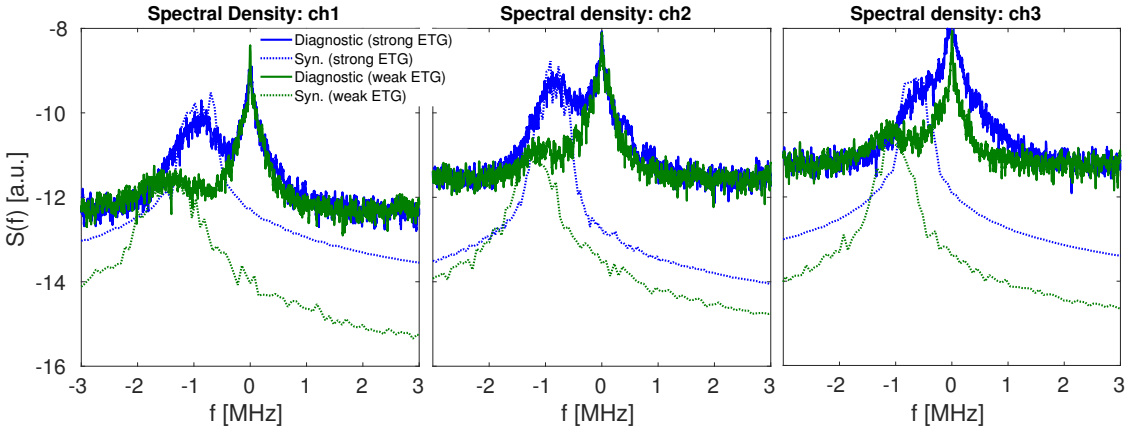

Figure 10: Frequency spectrum of high-k fluctuations from channels 1, 2 and 3 and synthetic frequency spectrum corresponding to flux-matched (P_e) simulations. Synthetic spectra includes Doppler shift. For the strong ETG case spectra are computed from the $\sigma(\nabla n), q, \hat{s}$-scan simulation (purple square in figure 7.a)), while the weak ETG spectra use the $\sigma(\nabla T, \nabla n)$-scan simulation (magenta dot in figure 8.b)). Experiment and simulation are plotted with the same color code, blue for the strong ETG condition and green for the weak ETG condition. The experimental f-spectra are normalized by the same constant (both the strong and weak ETG conditions), preserving the fluctuation level ratio (section VI for details). Putting aside the zero-frequency peak in the measurement and the different background noise levels, frequency spectra comparisons yield remarkable agreement for all channels. The predicted spectral peak lies within 5% of the measurement value for all channels, while the spectral width lies within 20% for the strong ETG condition (table 2), and yields a slight underprediction up to 40-50% for the weak ETG condition (table 3). This could be due to the low turbulence level and reduced signal to noise ratio possibly providing a non-reliable measure of the spectral width, but also possibly due to the discreteness in the simulated turbulence prediction for the spectral width, or some important physics processes not implemented in the simulation, such as cross-scale coupling effects.

Frequency spectra comparisons (Strong ETG)						
	ch 1		**ch 2**		**ch 3**	
/[MHz]	$< f >$	W_f	$< f >$	W_f	$< f >$	W_f
Exp.	-0.91	0.21	-0.81	0.19	-0.62	0.19
Sim.	-0.97	0.18	-0.76	0.16	-0.65	0.14

Table 2: Frequency spectra comparisons for the strong ETG condition corresponding to the frequency spectra in figure 10 (blue curves). The spectral peak $< f >$ and spectral width W_f are compared between experimental high-k measurements and synthetic predictions, shown for the three operating channels of the high-k diagnostic (in [MHz]). The uncertainty in $< f >$ and W_f is estimated to be $\sim$ 15-20% computed by varying the prescribed frequency window size. Remarkable agreement is obtained in the spectral peak $< f >$ for all channels, a confirmation of the Doppler shift dependence of the synthetic signal and of the correct wavenumber sampled by the high-k diagnostic (recall the Doppler shift frequency f_D is proportional to the measurement k). Good agreement is also found for the spectral width in this condition, lying within $\sim 20\%$ of the experimentally detected value for all channels.

be due to the reduced signal to noise ratio for the strong ETG condition, which could provide an unreliable measure of the experimental spectral width, or possibly due to the discreteness in the simulated turbulence producing a reduced spectral width prediction. Additionally, it could also be that the simulations presented here do not model an essential physics process that impacts the spectral width, such as, possibly, cross-scale interactions. In this work we were limited by computational resources and multiscale simulation was out of the scope.

Frequency spectra comparisons (Weak ETG)						
	ch 1		**ch 2**		**ch 3**	
/[MHz]	$< f >$	W_f	$< f >$	W_f	$< f >$	W_f
Exp.	-1.39	0.36	-1.21	0.34	-1.07	0.34
Sim.	-1.4	0.26	-1.23	0.22	-1.05	0.19

Table 3: Frequency spectra comparisons for the weak ETG condition corresponding to the frequency spectra in figure 10 (green curves). The spectral peak $< f >$ and spectral width W_f are compared between experimental high-k measurements and synthetic predictions, shown for the three operating channels of the high-k diagnostic (in [MHz]). The uncertainty in $< f >$ and W_f is estimated to be $\sim$ 20-25% computed by varying the prescribed frequency window size. Remarkable agreement is obtained in the spectral peak $< f >$ for all channels. However the spectral width tends to be slightly underpredicted by the synthetic diagnostic (up to 40-50%). This could be due to the reduced signal to noise ratio in this condition, possibly providing an unreliable measure of the spectral width. Alternatively it could be that the discreteness in the simulated turbulence or a missing physics process not being modeled are resulting in a reduced spectral width prediction.

One of the difficulties towards using $< f >$ and W_f to discriminate between simulations stems from the high sensitivity of these quantities to external factors such as Doppler shift and the diagnostic wavenumber resolution. In particular, the spectral peak $< f >$ is completely dominated by Doppler shift. Even with modest values of toroidal rotation ($M \sim 0.2$) and modest 'high-k' wavenumbers in the range $k_\theta \rho_s \sim$ 3-5, we find a Doppler shift frequency $f_D = \omega_0 k_\theta r/q \approx$ 1-1.5 MHz (where ω_0 is the toroidal rotation frequency and q the safety factor). This frequency is found to be at least a factor of 10 higher than the plasma-frame propagation frequency of fluctuations, on the order $< f >_{\omega_0=0} \approx$ 10-100 kHz. Any information about the plasma-frame propagation is essentially lost in the measurement. Experiment/model comparisons showing agreement in the frequency spectral peak will not necessarily imply a close agreement in the plasma-frame propagation frequency of fluctuations, but rather that Doppler shift is correctly implemented in the model. We observe that the experimental spectral peak can be consistently reproduced for essentially all turbulence models (as long as the correct Doppler shift value is applied).

With respect to the spectral width of fluctuations W_f, the effect of Doppler shift is less pronounced but can still modify W_f by factors of up to $\sim$ 2. Questioning whether the spectral width W_f could be a useful metric to discriminate against turbulent models, we compare the predicted values for W_f from the seven 'big-box' electron-scale simulations in the absence of Doppler shift (five strong ETG and two weak ETG 'big-box' electron-scale simulations), *ie*. we compare the predicted, intrinsic plasma frame value of the spectral width for different drive terms. In figure 11 are plotted the frequency spectra from the five strong ETG 'big-box' e- scale simulations in a), and in b) for the weak ETG case, corresponding to channel 1. The first thing to note is that in the absence of Doppler shift (*ie.* in the plasma-frame, setting the background toroidal rotation frequency $\omega_0 = 0$), the average spectral peak lies in the range $< f >_{\omega_0=0} \approx$ 10-100 kHz, clearly underpredicting the experimental values of $< f >^{\text{exp}} \approx$ 1-1.5 MHz, as previously mentioned (figure 2). More importantly, the spectral width W_f is shown to exhibit similar values in the plasma frame for all seven spectra, in the range $W_f \approx$ 150-180 kHz. These values are within the 1-σ uncertainty range $\sim$ 15-20%, determined via variations in the prescribed frequency band of analysis. Notice how W_f is insensitive to extremely different turbulence drive conditions, spanning up to five orders of magnitude in the total spectral power: from fully suppressed ETG in b) (red curve) to highly unstable ETG in a). We note how the same numerical resolution was employed for the strong and weak ETG simulations respectively, and the measurement wavenumber was similar in both conditions ($k_\theta \rho_s \sim -5.4$ was unchanged, while k_r changed by $\sim$ 30%). This suggests that the spectral width is not a good discriminator of turbulence models with varying turbulence drive terms. The spectral width measured by the high-k system could then be strongly influenced by the specific characteristics of the measurement, and only weakly by the intrinsic

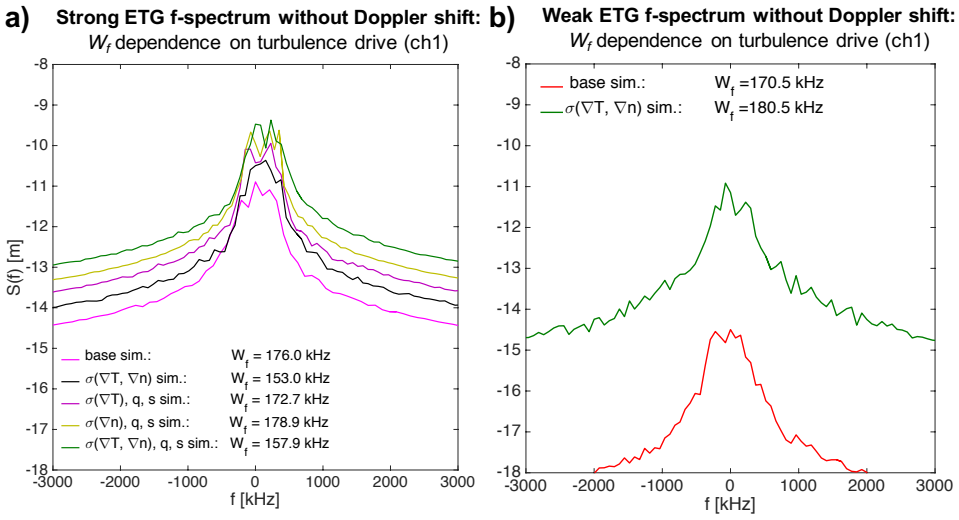

Figure 11: Synthetic frequency spectrum of high-k fluctuations from channel 1, respectively for five strong ETG 'big-box' electron-scale simulations in a) and 2 simulations for the weak ETG condition in b). The synthetic spectra was computed in the absence of Doppler shift ($\omega_0 = 0$), showing that the average plasma-frame propagation frequency of fluctuations $< f >_{\omega_0=0}$ is ~ 10 smaller than the experimentally measured value, suggesting Doppler shift completely dominates the frequency spectrum. Surprisingly, the spectral width W_f exhibits similar values for all 7 simulations, in spite of the fact that simulations span five orders of magnitude in spectral power. This suggests that the spectral width is not sensitive to the plasma turbulence conditions, but is mainly determined by the specific characteristics of the measurement (detected wavenumber $\vec{k}_+$ and diagnostic resolution Δk, implemented here via the synthetic diagnostic) and not by the intrinsic plasma-frame turbulence conditions. This figure shows how neither the spectral peak $< f >$ nor the spectral width W_f provide reliable metrics to discriminate against the different turbulent models with varying drive terms.

plasma-frame turbulence conditions. In fact, calculations presented in appendix C show that the predicted spectral width can be sensitive to the diagnostic resolution Δk in combination with Doppler shift when both values are large. This would mean that the high-k measurement can strongly prescribe the value of the measured spectral width. For the current experimental value of the wavenumber resolution and toroidal rotation frequency ω_0 we find that the lab-frame spectral width W_f lies within experimental uncertainty of the intrinsic plasma-frame value for the strong ETG case, but has a more important impact in the weak ETG case due to the enhanced toroidal rotation.

We can assess the influence of Doppler shift on W_f for the strong ETG condition by comparing the orange curve in figure $11.a)$ (without Doppler shift) to channel 1 from table 2 (with Doppler shift), corresponding to the same simulation. This shows that Doppler shift does not have a substantial impact on W_f in this condition, varying from $W_f^{\omega_0=0} \sim 178.9$ kHz to $W_f \sim 182.5$ kHz. However, a similar analysis performed for channel 1 of the weak ETG condition (comparing green curve from figure $11.b)$ to channel 1 in table 3) shows the spectral width varies from $W_f^{\omega_0=0} \sim 180.5$ to $W_f \sim 261.5$ kHz, lying clearly outside the uncertainty range. This change can now be exclusively attributed to Doppler shift: even though the measured wavenumber is unchanged between the strong and weak ETG conditions ($k_\theta \rho_s \sim -5.4$ for channel 1), the toroidal rotation frequency was $\sim 60\%$ higher for the weak ETG case, giving rise to a higher spectral peak $< f >$ and a widening of the spectral width W_f.

Together, the analysis performed on the synthetic spectra and the comparisons to experiment have shown that frequency spectra characteristics do not provide reliable validation metrics, as they are insensitive to the plasma frame turbulence conditions, and strongly affected by external factors such as Doppler shift (calculations shown in appendix C additionally show how the spectral width can be strongly prescribed by the high-k measurement). However, it is still highly valuable to perform experiment/model comparisons of the frequency spectrum. A successful validation activity should still be able to match the frequency spectra characteristics, independently of their ability to discriminate between models, and could be used as a test

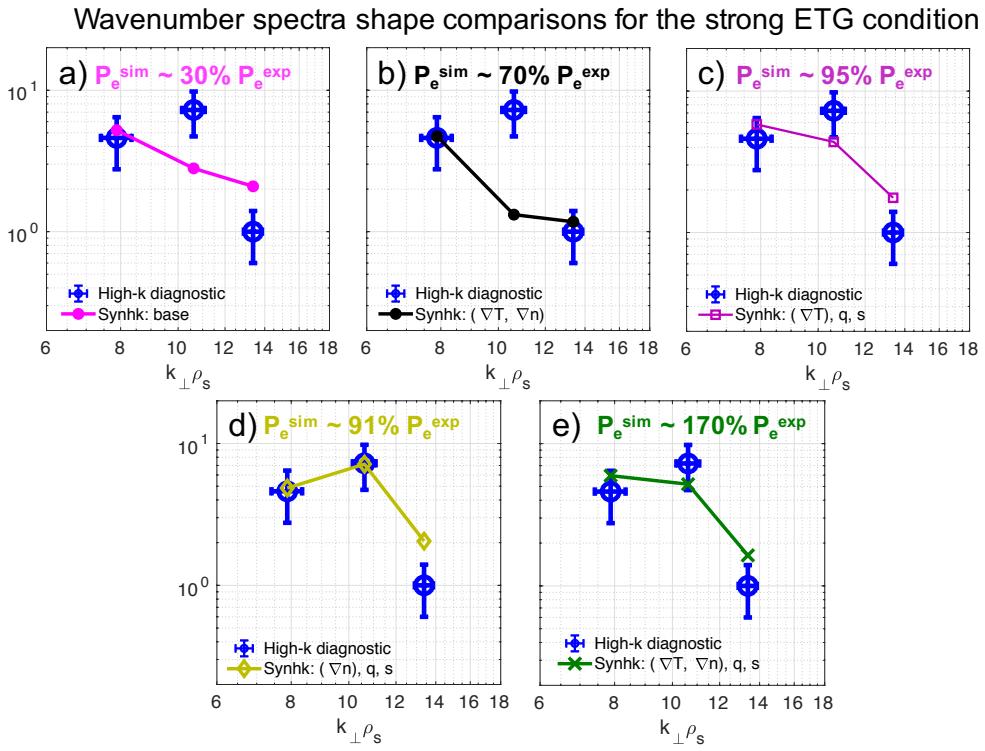

Figure 12: Wavenumber spectra shape comparisons for the strong ETG condition. Experimentally detected high-k fluctuations by the high-k system are shown by blue circles. Synthetic turbulence fluctuations correspond to the five 'big-box' electron-scale-scale simulations in figures 7.a) and b). The total synthetic fluctuation power is scaled (only in this plot) to minimize distance with respect to experimental fluctuation levels, allowing accurate comparisons of the wavenumber spectra shape. **a)** 'big-box' electron-scale simulation run for experimental base values of electron temperature and density gradient. **b)** Scaled 'big-box' electron-scale simulation by temperature and density gradient within 1-σ uncertainty. **c)** Flux-matched 'big-box' electron-scale simulation run with scaled temperature gradient within 1-σ uncertainty and an additional -10% in safety factor q and $+20\%$ in magnetic shear $\hat{s}$. **d)** Flux-matched 'big-box' electron-scale simulation run with scaled density gradient within 1-σ uncertainty and an additional -10% in safety factor q and $+20\%$ in magnetic shear $\hat{s}$. **e)** Scaled 'big-box' electron-scale simulation by temperature and density gradient within 1-σ uncertainty and an additional -10% in safety factor q and $+20\%$ in magnetic shear $\hat{s}$.

of the synthetic diagnostic. Overall, the frequency spectra comparisons are highly satisfactory, but suggest additional validation metrics are needed to accurately discriminate between simulations.

VI.b Wavenumber spectra comparisons

Here we use the k-spectrum shape and the fluctuation level ratio between the strong and weak ETG conditions as validation metrics to perform the comparisons to experiment. In this section we will compare the wavenumber spectrum shape from the five 'big-box' electron-scale simulations for the strong ETG condition. With respect to the fluctuation level ratio, we will compare the ratio between the five 'big-box' electron-scale simulations for the strong ETG case with respect to the flux-matching simulation for the weak ETG case (magenta dot in figure 8.b)). No use will be made of the 'big-box' e- scale base simulation for the weak ETG case (using experimental parameters, magenta dot in 8.a)).

Figure 12 shows the five k-spectra shape comparisons for the strong ETG condition. Only in this figure, the synthetic k-spectrum is scaled in order to minimize the corresponding distance to the experimental spectrum given by the blue circles - this yields a visual comparison of the **shape**. The ETG drive is increased from a) to e). The predicted P_e is shown for each simulation. Close inspection of figure 12 highlights that

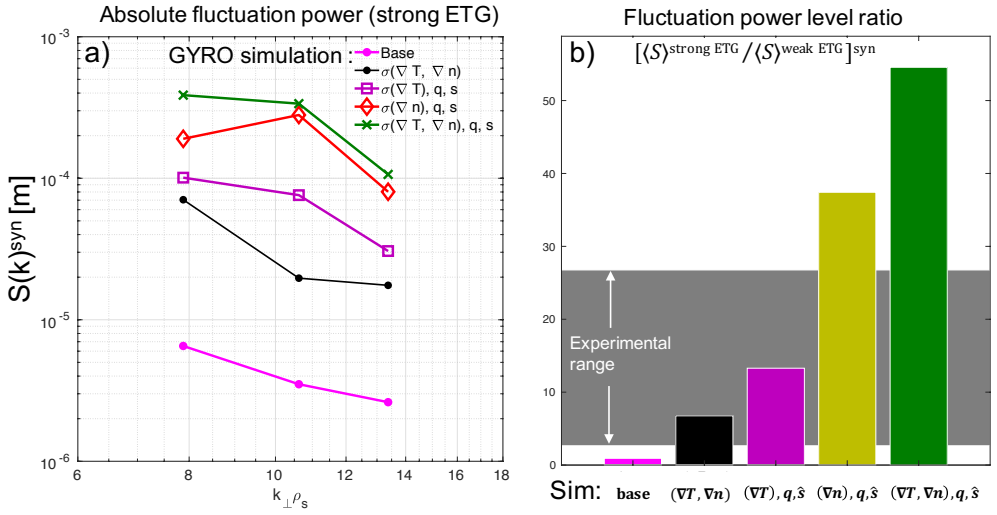

Figure 13: a) Absolute fluctuation power $S(k)^{\text{syn}}$ (wavenumber spectrum) from 'big-box' electron-scale-scale simulation carried for the strong ETG condition. Combinations of input drive terms correspond to the same simulations in figure 12 (here the synthetic k-spectrum has not been scaled but has absolute units). **b)** Synthetic fluctuation level ratio between strong ETG and weak ETG conditions. In computing the synthetic fluctuation level ratio $[\langle S\rangle^{\text{strong ETG}}/\langle S\rangle^{\text{weak ETG}}]^{\text{syn}}$, the strong ETG fluctuation power levels correspond to 'big-box' electron-scale simulations from a), while the weak ETG fluctuation power levels are computed from the same simulation for all comparisons in b), corresponding to the flux-matching weak ETG 'big-box' electron-scale simulation (magenta dot in figure 8). The gray band denotes the experimental fluctuation power level ratio between the strong and weak ETG conditions, with value of $\sim 14.7 \pm 80\%$. Simulations with scaled $(\nabla T, \nabla n)$ and $(\nabla T), q, \hat{s}$ lie within the 1-σ experimental range of the fluctuation level ratio.

simulations that underpredict the electron thermal power P_e are unable to match the shape of the spectrum: increasing the ETG drive from $a)$ to $b)$ does not improve the k-spectrum shape, but worsens it. Importantly, simulations with scaled q and shear produce a much better agreement, two of them being able to match P_e (c) and $d)$). This suggests q and $\hat{s}$ play a crucial role in determining the spectral shape in these conditions. The scale in the y-axis is logarithmic, meaning the shape of the wavenumber spectrum is a sensitive metric for comparisons. In particular, close to perfect agreement is found in $d)$. The spectral shape comparisons are quantified via a validation metric $\mathbf{d_{shape}}$ and shown in table 4. The impact of q and $\hat{s}$ on the shape of the spectrum is consistent with the current ramp-down performed in this plasma discharge, which was shown to modify the local values of q and $\hat{s}$. These comparisons are experimental evidence suggesting that the current ramp-down modified local values of q and shear, which had an impact on the shape of wavenumber spectrum of electron-scale density fluctuations.

The fluctuation level ratio comparisons are shown in figure 13. $13.a)$ shows the synthetic k-spectra from the five strong ETG simulations, this time plotted in absolute units (meters [m] due to normalization with respect to the toroidal length of the scattering volume, non-existent in the 2D implementation). We note how the spectral power at the measurement wavenumbers $\vec{k}_+$ is highly sensitive to the ETG drive, spanning about 2 orders of magnitude. This is highly desirable for a validation metric, allowing clear discrimination of the different turbulence models. Note how the predicted P_e is much less sensitive to the drive terms (spanning only a factor of ~ 5). In figure $13.b)$ are shown the comparisons of the fluctuation level ratio between the strong and weak ETG conditions. The average fluctuation power between the three channels is computed for each simulation in a). The average power for the strong ETG simulations in $13.a)$ is normalized by the average fluctuation power from the flux-matching, weak ETG simulation (magenta dot in figure $8.b)$). This yields a specific value of the fluctuation level ratio for each simulation in $a)$. This ratio can be readily compared to the experimental fluctuation ratio. Figure $13.b)$ shows that two simulations clearly lie within the 1-σ experimental range: in black with scaled $(\nabla T, \nabla n)$ and in purple with scaled $(\nabla T), q, \hat{s}$. Notice how the latter was a flux-matching simulation, also producing reasonable agreement in the wavenumber spectrum shape (figure $12.c)$). The sensitivity of the ratio with varying drive terms motivated the use of a

logarithmic metric, as explained in section IV. Note how simulations in purple and orange in 13.b) (scaled $(\nabla T), q, \hat{s}$ and $(\nabla n), q, \hat{s}$) were both clearly matching P_e and the spectral shape (figure 12), but they exhibit factors of 2-3 difference in fluctuation level: again a confirmation of the sensitivity and pertinence of the metric. Note additionally how the predicted P_e was higher for the purple case but the fluctuation level was lower than the orange case, consistent with the change in the shape of the wavenumber spectrum between the two cases. The fluctuation level ratio comparisons are also quantified via a validation metric $\mathbf{d_{ratio}}$, as shown in table 4.

VI.c Validation metrics to quantify overall simulation fidelity

Having identified three meaningful metrics to establish direct experiment/model comparisons (electron thermal power level $\mathbf{P_e}$, fluctuation level ratio $\mathbf{ratio}$ and wavenumber spectral shape $\mathbf{shape}$), we set out to quantify the overall fidelity of the different simulations. As described in section IV, each validation observable $\mathbf{X}$ is compared to experiment via a validation metric of distance $\mathbf{d_X}$, itself used to compute a normalized measure $\mathbf{R_X}$ between 0 and 1 for each comparison, where 0 is indicative of perfect agreement while 1 indicates complete disagreement. If one were to assume an experimental uncertainty equal to the simulation uncertainty, a distance $\mathbf{d_X}$ between experiment and simulation of one sigma would correspond to a value of $\mathbf{R_X} \approx 0.04$. A distance $\mathbf{d_X}$ of two sigma would correspond to $\mathbf{R_X} \approx 0.5$, while three sigma would correspond to $\mathbf{R_X} \approx 0.9$.

Table 4 summarizes these most pertinent comparisons from sections V and VI. In this table, the simulations are shown with increasing ETG drive as in figure 12. Focusing on the third column, the bounded metric associated to electron thermal power $\mathbf{R_{Pe}}$ shows a clear disagreement for the base case simulation as expected ($\mathbf{R_{Pe}} \approx 0.999$). As we increase ETG drive, the P_e comparisons improve and $\mathbf{R_{Pe}}$ decreases down to 0.006 for the simulation with scaled $\sigma(\nabla T), q, \hat{s}$. The comparison worsens again as we further increase ETG drive in the last row for scaled $\sigma(\nabla T, \nabla n), q, \hat{s}$, which predicted $\sim 170\%$ of the experimental P_e level. A similar analysis can be performed for the 5^{th} column of the fluctuation level ratio. With respect to the shape of the k-spectrum, the best agreement is found for the simulation with scaled $\sigma(\nabla n), q, \hat{s}$, as suggested by inspection of figure 12.

	Validation table						
	$\mathbf{P_e}$		**Fluct. level ratio**		k**-spec. shape**		**Composite metric**
Simulation	$\mathbf{d_{Pe}}$	$\mathbf{R_{Pe}}$	$\mathbf{d_{ratio}}$	$\mathbf{R_{ratio}}$	$\mathbf{d_{shape}}$	$\mathbf{R_{shape}}$	$\mathbf{M_1}$
Base	3.22	0.999	1.11	0.18	5.08	1	0.672
$\sigma(\nabla T, \nabla n)$	1.27	0.29	0.74	0.046	4.24	0.99	0.472
$\sigma(\nabla T), q, \hat{s}$	0.215	0.006	0.38	0.01	2.52	0.98	0.397
$\sigma(\nabla n), q, \hat{s}$	0.397	0.012	1.47	0.47	1.53	0.53	0.402
$\sigma(\nabla T, \nabla n), q, \hat{s}$	3.12	0.998	1.74	0.72	1.78	0.76	0.762

Table 4: Validation table with the most relevant metrics allowing accurate discrimination between models: the electron thermal power $\mathbf{P_e}$, the electron-scale fluctuation level ratio and the k-spectrum shape. Comparisons are only made for the five 'big-box' electron-scale-scale simulations in the strong ETG case in figure 7. To compute the fluctuation level ratio between the strong and weak ETG conditions, we used the flux-matching 'big-box' electron-scale-scale simulation for the weak ETG case (magenta dot in figure 8). The parameters $\mathbf{d_X}$ denote a validation 'distance' for the comparison of each observable $\mathbf{X}$, while $\mathbf{R_X}$ denotes the corresponding bounded error metric between 0 and 1 and computed from $\mathbf{d_X}$. The composite metric $\mathbf{M_1}$ quantifies the overall model agreement, with 0 indicating perfect agreement and 1 indicating complete disagreement. Definitions for $\mathbf{d_X}$, $\mathbf{R_X}$ and $\mathbf{M_1}$ are taken from [33, 34].

A composite metric $\mathbf{M_1}$ combining the different metric results is used to quantify the overall fidelity of each model. Here $\mathbf{M_1} = \sum_i h_i \mathbf{R_i} / h_i$, where i is summing over the different observable comparisons ($\mathbf{Pe}, \mathbf{ratio}, \mathbf{shape}$) and h_i is the weight given to each comparison according to its level in the validation hierarchy [30]. Here we chose $h_i = 0.5$ for the $\mathbf{Pe}$ comparisons, while $h_i = 1$ for the $\mathbf{ratio}$ and $\mathbf{shape}$ comparisons of the wavenumber spectrum (a discussion on the specific values of h_i is provided in the next section).

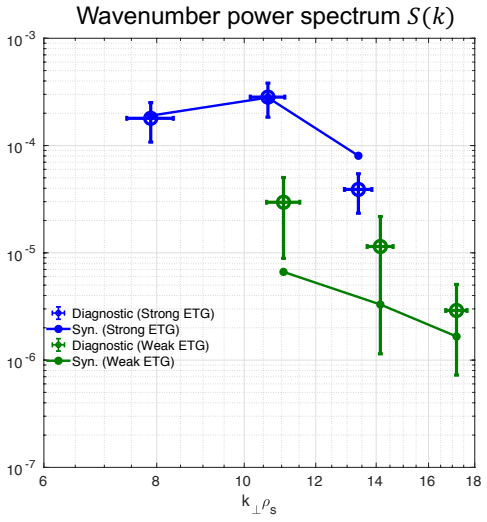

Figure 14: Wavenumber spectra comparisons between experimental spectra and synthetically generated spectra from flux-matched (P_e) simulations, both for the strong and weak ETG conditions. The strong ETG synthetic spectra are generated from the $\sigma(\nabla n), q, \hat{s}$-scan simulation that predicts $P_e^{sim} \approx 95\%\ P_e^{exp}$, and correspond to the frequency spectra in figure 10, also producing a best match in the validation metric comparisons ($cf.$ table 4). The experimental strong ETG k-spectrum is scaled in order to minimize the 'distance' with the synthetic k-spectra. The weak ETG experimental k-spectra are scaled by the same constant in order to preserve the fluctuation level ratio. The weak ETG simulation spectrum corresponds to a 'big-box' electron-scale simulation that matches P_e within experimental uncertainty ($P_e^{sim} \approx 80\%\ P_e^{exp}$). These results give strong evidence suggesting ETG fluctuations alone may be the main mechanism driving electron thermal transport in both the strong and weak ETG conditions.

The corresponding values of $\mathbf{M_1}$ for each simulation are shown in the last column of table 4, with 0 indicating agreement and 1 indicating disagreement. Counter-intuitively, the overall comparison is worse for the last simulation with scaled $\sigma(\nabla T, \nabla n), q, \hat{s}$ than for the base case, owing to the overprediction of P_e and especially to the overprediction of the fluctuation level ratio. The best overall agreement is found for simulations with scaled $\sigma(\nabla T), q, \hat{s}$ and $\sigma(\nabla n), q, \hat{s}$, yielding an approximate value of ~ 0.4. Here a value of $\mathbf{M_1}$ of 0.4 indicates an agreement between experiment and simulation roughly between one and two sigma uncertainty (if one were to assume identical simulation and experimental uncertainty). It is satisfactory that both simulations are able to match the experimental level of electron thermal power. The wavenumber power spectrum from experiment and simulation for one of the two best matched cases is shown in figure 14.

VII Caveats of this work

Nonlinear gyrokinetic simulations

A substantial effort has been made in this work to include as many relevant physics processes as was reasonably possible within our computational constraints and availability of physics models implemented in the simulations: resolving ion to electron-scales with fully electromagnetic fluctuations, 3 kinetic species, background parallel and toroidal flow shear, etc. However, there is no question that there are missing physics processes not being directly modeled within this work that could potentially impact our conclusions. We have already discussed the potential importance of cross-scale interactions between electron-scale and ion-scale turbulence as well as profile variation, none of which are modeled in this work. Concerning profile variation, past work performed for the weak ETG condition showed that global gyrokinetic simulation by the GTS code [39] in fact predicts negligible electron thermal transport from ion-scale turbulence, similarly to what we observed when using the experimental profile values as input. Additionally, nonlinear GYRO

simulations do not include a fast ion population, which in this case can be justified via the small fast ion pressure with respect to the electron pressure, on the order to 10%. Negligible impact from fast ion population is expected.

It is worth saying a few words with respect to the collisional model and high collisionality conditions in these plasmas ($\nu_{ei} \sim 1$ c_s/a), as well as the implementation in GYRO. Electron collisions were included in these simulations via a pitch angle scattering collision operator. Energy diffusion is not modeled in the collision operator, which could be relevant in these high collisionality conditions. Improved collisional models such as implemented in CGYRO [40] could be needed. Ion-ion collisions have also not been modeled in the present work. Although ion-ion collisions generally play a negligible effect on the predicted transport levels, they may become important in near marginal conditions and high collisionalities such as we observe for the weak ETG condition, as was suggested by past work [41–44]. In fact, recent simulation work on these NSTX H-mode plasma conditions (not presented here) suggests ion-ion collisions could play an important role in determining the saturation of the zonal flow amplitude in conditions close to the nonlinear threshold. Ion-ion collisions are shown to produce an additional damping mechanism on the zonal flow, which is weakened in near marginal conditions. This results in an increase of the ion-scale turbulence level when close to the nonlinear threshold ($a/L_{Te} \sim 5$ in figures 8, 9), but not for far from threshold conditions. Overall, including ion-ion collision was shown to bring the effective nonlinear threshold of ion-scale turbulence closer to the experimentally detected base profile values, while still predicting a factor of 10 larger values than the experimental levels within experimental uncertainty in $\nabla T, \nabla n$. Ion-ion collisions were shown to have small impact on electron-scale turbulence simulation. However, since negligible thermal transport can be expected on the ion channel (owing to the neoclassical P_i constraint), these simulations point to the same result as we found here: stable ion-scale turbulence in the TEM-stable regime is the most experimentally relevant, and pointing to ETG as the dominant mechanism to explain P_e at the weak ETG condition.

In order to test the impact that a further improved collisional model could have on our conclusions, linear simulations with the CGYRO code were performed for the base case weak ETG condition, for which GYRO predicted stable ion-scale turbulence. Preliminary linear CGYRO simulations showed that the linear growth rates lied within 25% of the GYRO predictions. Corresponding ion-scale nonlinear CGYRO simulation showed ion-scale turbulence was suppressed by strong $E \times B$ shear, an identical prediction to the one made by GYRO. These tests having been performed, we have confidence to assert that the conclusions that one can draw by using improved collisional models will remain strictly the same as the ones presented here: near threshold behavior of ion-scale turbulence in the TEM-stable regime, while electron-scale turbulence is the dominant heat loss mechanism and cross-scale interactions possibly important.

One open question concerns the extremely high transport predictions by ion-scale simulation for the weak ETG condition, observed when $\nabla T, \nabla n$ were scaled by 1-σ. Ion-scale simulations not presented here show that the unstable TEM-branch was extremely sensitive to the inclusion of compressional magnetic field perturbations $\delta B_{||}$, exhibiting almost complete turbulence suppression when $\delta B_{||}$ perturbations were not included in GYRO. In addition, linear GYRO simulations scanning the electron beta β_e showed how the experimental beta value sits within $\sim 20\%$ of the KBM threshold. This suggests that KBM could have a role to play in this condition. Previous work performed for an NSTX H-mode [1] has in fact highlighted a hybrid TEM/KBM mode can produce experimentally relevant P_e. Further analysis would need to be performed to shed further light on this issue.

Synthetic turbulence comparisons

The direct frequency and wavenumber spectra comparisons performed as part of this study have not only provided incredibly valuable constraints to understand the turbulent origins of anomalous electron thermal transport, but have also yielded a profound understanding of the high-k measurement at NSTX. Frequency spectra comparisons have confirmed that Doppler shift dominates the measured frequency spectrum of fluctuations. Past validation studies [30] have indicated that "comparison of time series data provide some of the most fundamental validation tests". However, we have learned in this validation work that it is nearly hopeless to discern between turbulence models by direct comparisons of the frequency spectra. The

spectral peak $< f >$ has shown to be dominated by Doppler shift, while the spectral width W_f is insensitive to the state of the turbulence. Appendix C shows how the spectral width can be strongly affected by the wavenumber resolution of the diagnostic. This means the spectral width W_f measured by the high-k system can be strongly constrained by the measurement characteristics, and only weakly dependent on the intrinsic turbulence characteristics. Together these arguments suggest that the measured frequency spectrum does not provide strict constraining metrics for the discrimination of turbulence models.

There is no doubt that the synthetic high-k turbulence predictions presented here suffer from uncertainties and inaccuracies emanating from the approximations made in the synthetic model. One important approximation lies in the $2D$ implementation. This approximation is justified by the toroidal localization of the high-k system as detailed in [31, 32]. Although not shown here, work recently performed in the context of this paper has suggested that 3D effects are expected to be negligible for the high-k scattering diagnostic, as long as the scattering condition $\vec{k} \cdot \vec{B}$ is satisfied. The inclusion of a finite toroidal extent in the scattering volume and the exploration of 3D effects on turbulence scattering measurements could be the object of a future publication.

An additional approximation in the synthetic diagnostic described here is based on a constant-k: the same turbulence wavenumber is sampled within the whole simulation domain. In fact, the measurement wavenumber provided by ray-tracing calculations is only representative of the central ray of the input microwave beam (which is the one that contributes the most to the synthetic signal), but a slightly different wavenumber is sampled by the diagnostic at different radial, poloidal and toroidal locations within the scattering volume. To assess the impact of this constant-k approximation, additional ray tracing calculations were performed for non-central rays within the scattering volume, showing that the measurement-k can at most vary $\sim 20\%$ within the scattering volume. This would have a small impact on the synthetic frequency and k-spectra characteristics when compared to other factors such as the simulation wavenumber resolution or Doppler shift, especially since the non-central rays which are associated to a different sampled k are filtered down with respect to the central ray, and produce smaller power contributions to the synthetic signal. However, it is possible that taking into account the spatial variation of k within the scattering volume could recover the underpredictions observed in the frequency spectral width in tables 2, 3.

All simulations presented here are based on the local approximation, representing only one flux surface, and background profiles are linear in the simulation domain. However it can be seen in figure 1 that the scattering volume has a finite radial extent, meaning the high-k measurement has contributions from turbulence fluctuations present at different flux surfaces and subject to different drive terms. However here once more we can argue that it is the central flux surface containing the scattering location that will have the biggest contribution to the high-k spectra while turbulence fluctuation contributions from other nearby flux surfaces are filtered down by the filter and will likely have negligible contributions to the high-k signal. As can be seen from figure 1, the background temperature and density gradients have small profile variations within the scattering volume, which are in fact captured via the sensitivity scans in the background temperature and density gradients presented here. Global gyrokinetic simulation including profile variation would be able to give a more definite answer to confirm this hypothesis, however this was out of the scope of this work.

Within the realm of these approximations, we have reasonable confidence in the synthetic calculations presented. Additionally, as part of this work two synthetic diagnostics for high-k scattering were implemented: one based on filtering in real space, which is presented here, and another one based on filtering in wavenumber space. These two synthetic diagnostic implementations, which are analytically equivalent, also produce the same synthetic spectra within small percentage variations. This agreement further strengthens the confidence in our results.

Validation metrics

The use of validation metrics in this work has proven useful to quantify the experiment-model comparisons, as well as to discriminate between those that best match all the experimental measurements. However there

is no unique way to quantify these comparisons, and one could have opted to employ a different set of metrics. For example, the particular choice of the parameters $d_0 = 1.5$ and $\lambda = 0.5$ employed in the definition of $\mathbf{d_X}$ (equation 3, suggested in [33]) appear to be arbitrary. Here we try to provide an intuition behind the specific values of $\mathbf{d_X}$ shown. In the case of equal values of the experimental and simulation uncertainties (noted σ here), a value of $\mathbf{d_X} \approx 0.7$ maps to $\mathbf{R_X} \approx 0.04$, which roughly corresponds to a discrepancy between experiment and simulation of 1-σ. A value of $\mathbf{d_X} \approx 1.5$ maps to $\mathbf{R_X} \approx 0.5$, which roughly corresponds to a discrepancy of 2-σ. Keeping these values in mind, although approximative, can give an intuition for interpreting the numerical values of table 4.

The specific values of the weighting factors h_i employed in the computation of the composite metric $\mathbf{M_1}$ might also appear arbitrary. These should be chosen according to how the user would like to preferentially weigh each comparison with respect to others in a composite metric. In this work, we have chosen them to be an increasing function of the hierarchy level of each comparison (e.g. frequency spectra, wavenumber spectra, transport, etc.). This choice preferentially values the comparisons in the lower hierarchy levels, as suggested by [29, 30, 34]. The rationale behind this being the fact that comparisons at the lowest levels in the hierarchy are the ones that contain the most fundamental characteristics of the turbulence (frequency spectra, wavenumber spectra), and should set the most stringent conditions to match experiment.

In this work we have performed comparisons at three different levels in the validation hierarchy: frequency spectra comparisons (lowest level) via $< f >$ and W_f, wavenumber spectra comparisons (second level) via $\mathbf{ratio}$ and $\mathbf{shape}$, and thermal power comparisons (highest level) via $\mathbf{P_e}$. As was discussed in the text, comparisons of frequency spectra quantities are not useful for the discrimination of the numerical simulations, and are not included in the computation of the composite metric $\mathbf{M_1}$. We have settled to employ the weighting factors $h_i = 1$ for $\mathbf{ratio}$ and $\mathbf{shape}$, and $h_i = 0.5$ for $\mathbf{P_e}$. Although not shown here, the use of different weighting factors h_i, or even different composite metric definitions produces very similar final values. Most importantly, the order in the ranking of each of the five simulations compared to experiment in this work was shown to be unaltered by the use of different weighting factors and different composite metrics (we computed metrics $\mathbf{M_0}$ and $\mathbf{M_2}$ appearing in equations 36 and 38 from reference [34]). The ranking indicated by the last column of Table 4 appears to be a robust result.

VIII Conclusions and next steps

The extensive validation effort performed for this modest-beta NSTX NBI heated H-mode plasma has tried to shed some light towards understanding the origins of electron thermal transport in the core gradient region of spherical tokamak H-modes. By using a synthetic diagnostic for high-k scattering, we have shown the first simultaneous agreement between experiment and simulation of electron thermal transport, electron scale frequency spectra, the shape of the wavenumber spectra and fluctuation level ratio between strongly driven and weakly driven ETG turbulence conditions. This is the strongest experimental evidence to date that ETG turbulence can dominate electron thermal transport in the outer-core of modest-beta NSTX NBI-heated H-modes.

We have made extensive use of high-k scattering fluctuation data to place strict constraints on electron-scale gyrokinetic simulations. However, no specific experimental measurement of ion-scale turbulence nor zonal flows are used in this study, leaving them fully unconstrained. Additional fluctuation measurements would be needed to constrain turbulence fluctuations at ion-scales. In fact, our work has highlighted the possibility that ion-scale fluctuations could produce experimentally relevant transport levels, specifically for the weak ETG condition. Low-k fluctuation measurements such as provided by beam emission spectroscopy (BES), as well as fluctuation scattering diagnostics sensitive to low-k and intermediate range wavenumbers like Doppler back-scattering (DBS) and reflectometry would provide invaluable information to constrain our turbulence models at all relevant scales characteristic of microturbulence fluctuations. Zonal flow fluctuation measurements could also provide critical constraints, especially for the zonal flow dominated conditions at weak ETG exhibiting large zonal flow fluctuation amplitudes with respect to the turbulence fluctuation amplitudes predicted by GYRO. Additionally, measurements of internal magnetic field fluctuations via

cross polarization spectroscopy (CPS) would prove extremely valuable to diagnose the microtearing and KBM modes destabilized at high-β.

Several arguments presented here have pointed to multiscale simulation as a possible future avenue emanating from this work. To our knowledge no successful attempt has been made to carry out multiscale simulation of spherical tokamak plasma. Past work performed for C-Mod L-mode discharge [26, 27] showed how cross-scale coupling became important when ion-scale turbulence was near marginal - similarly to the weak ETG condition analyzed here. Operating at different parametric regimes, typically at higher β, increased shaping and higher toroidal rotation than standard tokamaks, multiscale simulation of spherical tokamak differs from current multiscale simulation efforts. Coupling multiscale simulation with fluctuation measurements and synthetic diagnostics would be critical towards determining the conditions required for multiscale simulation and the subsequent contributions to electron thermal transport in the ST. Additionally, owing to the large normalized value of the ion gyro-radius $\rho_\star$, it could be that global, multi-scale simulation including profile variation is required in an ST, representing even a greater computational challenge.

If the ultimate goal is to predict electron thermal transport and consequently the equilibrium plasma profiles of future high performance plasmas and fusion reactors, further emphasis should be placed on predictive transport simulations coupled with reduced transport models such as TGLF. This work has shown both the value as well as the cost of performing first-principles gyrokinetic simulations. Reduced transport models are, by nature, much less computationally intensive. Developing reduced transport models accurate to the high beta, high flow and low-aspect ratio ST regime should be a main research priority of future confinement studies in the ST. Additionally, a natural next step would be to couple TGLF to a synthetic high-k diagnostic, guiding the development of more sophisticated gyrokinetic simulations. Together, these efforts would drive the prediction and optimization of the performance of future fusion reactors.

A Experimental plasma parameters input in GYRO

Parameter	Strong ETG value	Weak ETG value
r/a	0.70839	0.67864
$R_0(r)/a$	1.5227	1.5922
$\Delta = dR_0(r)/dr$	-0.30412	-0.355
κ	2.1127	1.9792
$s_\kappa = rd\ln(\kappa)/dr$	0.153963	0.190789
δ	0.248295	0.167812
$s_\delta = rd\delta/dr$	0.323972	0.323851
q	3.7892	3.0723
$\hat{s}$	1.8047	2.3464
$\rho^\star$	0.00328	0.00382
$M = R\omega_0/c_s$	0.16	0.21
$\gamma_E = -r/qd\omega_0/dr$	0.1258	0.1646
$\gamma_p = -R_0 d\omega/dr$	1.0362	1.1558
ν_{ei}	1.379113	1.0310
Z_{eff}	1.9504	1.8425
n_D/n_e	0.78503	0.80371
n_C/n_e	0.035828	0.032715
T_D/T_e	1.1352	1.3964
T_C/T_e	1.1352	1.3964
a/L_{nD}	1.4973	4.0319
a/L_{nC}	-0.8769	4.0801
a/L_{ne}	1.0048	4.0576
a/L_{TD}	2.9599	3.0929
a/L_{TC}	2.9599	3.0929
a/L_{Te}	3.3626	4.5128
β_e	0.002736	0.002956
λ_D/a	0.0000371	0.0000417
$\mu = (M_D/m_e)^{1/2}$	60	60
$n_e\ [10^{19}/m^3]$	4.2714	3.4282
T_e [keV]	0.39060	0.40116
B_{unit}	1.44613	1.26986
$a\ [m]$	0.6012	0.5960
$c_s/a\ [10^5/s]$	2.2749	2.3252

Table 5: Plasma parameters input in GYRO simulations presented in this work.

B Numerical resolution tests

The nonlinear simulation transport predictions presented with the GYRO code have been extensively tested for convergence of the numerical solution. All numerical simulations were fully electromagnetic including $(\delta\phi, A_{||}, B_{||})$ field fluctuations, electron collisions ($\nu_{ei} \sim 1$), parallel flow and flow shear $(M, \gamma_E, \gamma_p) \sim$ (0.2-0.3, $\sim$ 0.1-0.2, ~ 1), and simulating 3 kinetic species (all species gyrokinetic for electron and 'big-box' electron-scale simulations, and using drift-kinetic electrons for ion-scale simulations). Ion-ion collisions ν_{ii} were not included in this study ($cf.$ conclusion VIII for additional details). All nonlinear simulations carried out as part of this work had the same poloidal resolution of 14 poloidal grid points ($\times 2$ signs of parallel velocity), 12 energies and 12 pitch-angles (6 passing + 6 trapped), which were tested for numerical convergence (although not shown here). The choice of numerical grids was consistent to previous convergence and accuracy tests for the GYRO code simulating micro-instabilities in the core of NSTX [1]. In this section are shown resolution and convergence tests carried out for electron and 'big-box' electron-scale

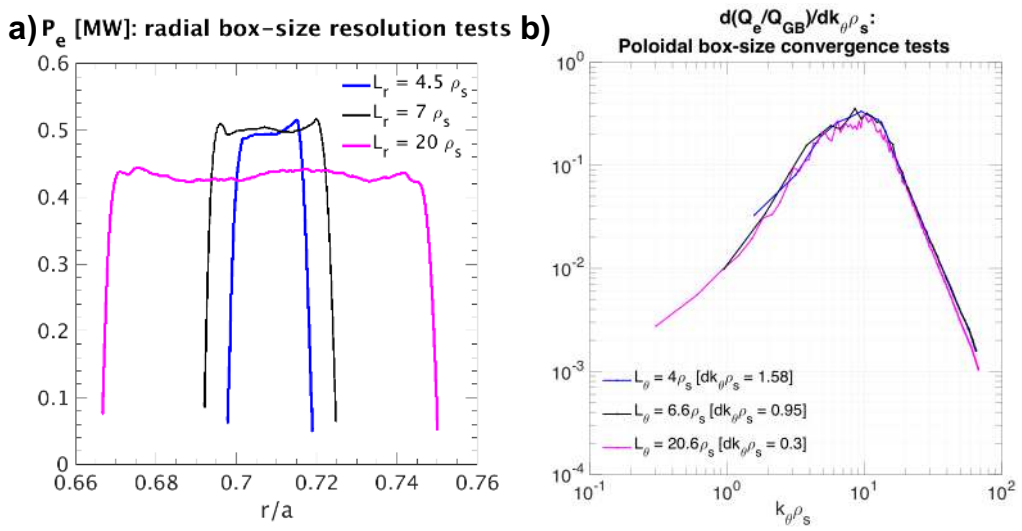

Figure 15: Radial and poloidal box-size convergence tests carried out for three different radial and poloidal box-size domains, corresponding to electron and 'big-box' electron-scale simulations for the strong ETG condition (experimental profile parameters were used as input). **a)** Radial box-size convergence tests: electron thermal power P_e is plotted versus normalized minor radius r/a for standard electron-scale simulations ($L_r = 4.5, 7\rho_s$) and 'big-box' electron-scale scale simulation ($L_r = 20\rho_s$). Similar electron thermal power is predicted for simulations with $L_r = 4.5\rho_s$ and $7\rho_s$, and a slight decrease of $\sim 10-15\%$ is observed when the radial box size is increased to $L_r = 20\rho_s$. Although this small decrease might be related to the impact of ion-scale modes associated to the larger box-size, or the shorter simulation time in the larger box-size due to limited computational resources ($T \sim 20$ a/c$_s$), we note however that this small difference lies well inside the simulation standard deviation, indicating convergence in radial box-size (note the error bars in the nearby blue and magenta dots of figure $7.a$) corresponding to the black and red curves in this plot). The GYRO buffer regions are not included in the calculation of the radial box-size. **b)** Poloidal box-size convergence tests: the fractional electron heat flux dQ_e per poloidal wavenumber increment d$k_\theta\rho_s$ is plotted versus normalized poloidal wavenumber $k_\theta\rho_s$ for standard electron-scale simulations ($L_\theta = 4, 6.6\rho_s$) and 'big-box' electron-scale scale simulation ($L_\theta = 20.6\rho_s$). A larger box-size is inversely proportional to the poloidal wavenumber increment d$k_\theta\rho_s$. All simulations resolve up to $k_\theta\rho_s^{max} \sim 65$. A similar k_θ spectrum is obtained for all simulations, in addition to similar predicted electron thermal power ($10-15\%$ agreement lying within the numerical standard deviation). We conclude convergence in poloidal box-size. The blue and magenta curves correspond to the nearby blue and magenta simulation points in figure $7.a$). The 'big-box' electron-scale scale simulation with the larger box-size consumed over $1M$ CPU h for completion NERSC's Edison supercomputer.

scale simulations for the strong ETG condition. Numerical convergence is shown in the radial box-size (L_r) and poloidal box-size domains (L_θ, inversely proportional to the poloidal wavenumber grid step d$k_\theta\rho_s$), in radial resolution for varying radial grid steps (dr/ρ_e) and in the maximum resolved poloidal wavenumber max$(k_\theta\rho_s)$.

Similar resolution and convergence tests were carried out for the weak ETG condition (not shown here). For this condition, it is worth separating the discussion on numerical convergence between the TEM-stable regime (figure $8.a$), and $8.b$) for $a/L_{Te} < 5$) and the TEM-unstable regime (figure $8.b$) for $a/L_{Te} > 5$). In the TEM-unstable regime, a slight disagreement between standard and 'big-box' electron-scale simulation is obtained in the heat transport predictions (for varying box-sizes (L_r, L_θ)), lying slightly outside the numerical standard deviation in predicted thermal transport ($cf.$ the rightmost blue and magenta dots from figure $8.b$)). These slight disagreements might be related to the highly unstable ion-scale modes present due to a strongly driven TEM, since some of those ion-scale modes were shown to play a small but non-negligible role in the larger box 'big-box' electron-scale-scale simulation for the weak ETG condition. We cannot at this point ensure numerical box-size convergence for the electron and 'big-box' electron-scale scale simulations presented for the weak ETG condition in the TEM-unstable regime. However, we have argued in this article how the TEM-unstable regime itself is not a good approximation of reality, since it cannot comply with the electron and ion thermal transport constraints (figure 9 for $a/L_{Te} > 5$),

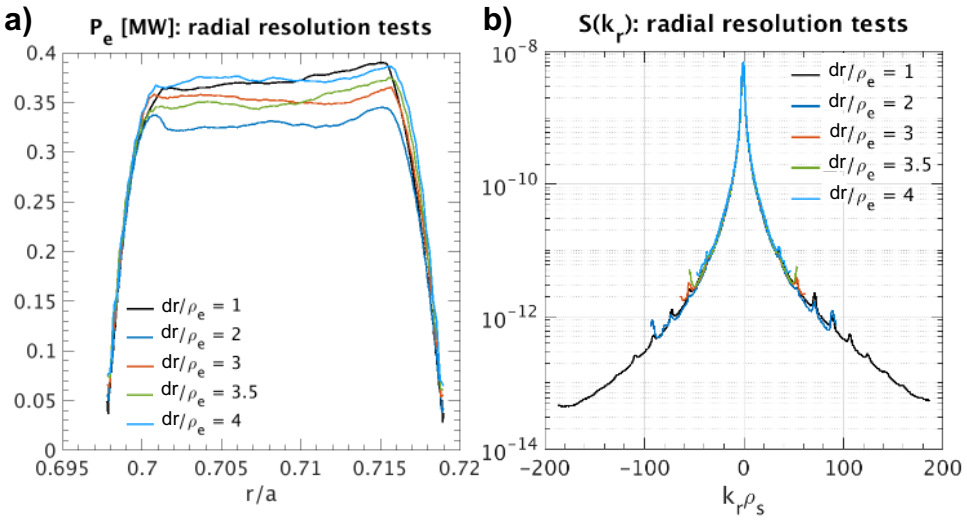

Figure 16: Radial resolution tests for five different radial grid steps $\mathrm{d}r/\rho_e$ of electron scale simulations for the strong ETG condition (experimental profile parameters were used as input). **a)** P_e[MW] is shown over the full radial domain, showing close P_e agreement between all simulation predictions. **b)** Radial wavenumber power spectrum associated to potential fluctuations $|\delta\phi|^2$. The field $\delta\phi$ exhibits the largest amplitude among all electromagnetic field components $(\delta\phi, \delta A_{||}, \delta B_{||})$, and is responsible for the dominant contribution to electron heat transport P_e by some ~ 2 orders of magnitude over contributions from shear $\delta A_{||}$ and compressional perturbations $\delta B_{||}$. All simulation resolution scans presented here are carried out on standard electron scale simulations with box-sizes $(L_r, L_\theta) = (4.5, 4)\rho_s$.

making the absence of absolute numerical convergence in the TEM-unstable regime a less relevant result. However, excellent numerical convergence is achieved between electron and 'big-box' electron-scale scale simulation in the TEM-stable regime. In fact, figure $8.a)$ shows identical heat transport predictions between electron and 'big-box' electron-scale simulation (overlaid blue and magenta dot in figure $8.a)$). In addition, good convergence was shown in radial resolution $\mathrm{d}r/\rho_e \sim 3$ and maximum resolved poloidal wavenumber $\max(k_\theta\rho_s) \sim 87$ for the weak ETG condition, both in the TEM-stable and TEM-unstable regimes.

Preliminary convergency tests were also carried out for ion-scale nonlinear simulations (specifically for the TEM-unstable branch in the weak ETG condition, which predicted $\sim \times 10$ the experimental electron thermal power level). These resolution tests for the ion-scale simulations confirmed the presence of the highly unstable TEM branch in the weak ETG condition, and predicted similar heat transport levels lying within the numerical standard deviation, both in the TEM-stable and unstable regimes.

With respect to the width of the damping regions employed by GYRO, they were chosen according to previous analysis and convergence tests performed for NSTX plasmas [1,11]. For ion-scale simulations, we used $\Delta_b = 8\rho_s$ as suggested in reference [1], while for electron-scale simulations $\Delta_b = 1\rho_s$ as suggested in reference [11]. For the 'big-box' electron-scale simulations we used $\Delta_b = 2\rho_s$. The widths for the 'big-box' electron-scale simulations were not explicitly tested for numerical convergence in the present study due to the large computational requirements of the simulations. As was shown in [11], the transport from standard electron-scale simulations resolving ETG turbulence was insensitive to buffer width values larger than $1\rho_s$. Since 'big-box' electron-scale simulations are only capable of resolving ETG turbulence, this would suggest a buffer width value of $\Delta_b = 2\rho_s$ is probably well converged and insensitive to further increases in Δ_b. However it is still possible that different buffer widths might affect the 'big-box' electron-scale simulation results. The fact that the same buffer width value was employed for all 'big-box' electron-scale simulations leads us to suspect this effect would be similar for all simulations. This could result in different predicted values of $\mathbf{P_e}$, $\mathbf{ratio}$ and $\mathbf{shape}$ (table 4), but would likely not affect the relative ranking of each simulation according to their degree of agreement with experiment, quantified by the composite metric.

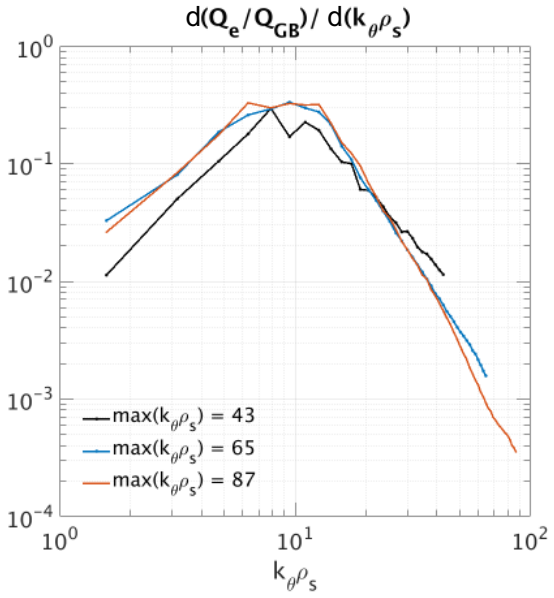

Figure 17: Poloidal resolution tests for 3 different values of the maximum $k_\theta \rho_s$ resolved in electron scale simulations for the strong ETG condition (experimental profile parameters were used as input). Simulation resolving a maximum $\max(k_\theta \rho_s) = 43$ exhibits reduced total electron thermal power (up to 30% reduction) and noticeably different spectral slope at high-k with respect to simulations with $\max(k_\theta \rho_s) = 65, 87$. Note however how negligible differences can be observed between simulations resolving $\max(k_\theta \rho_s) = 65$ $vs.$ $\max(k_\theta \rho_s) = 87$. This last point gives us confidence that simulation is converged for $\max(k_\theta \rho_s) = 65$. Consequently 'big-box' electron-scale scale simulations for the strong ETG case will resolve a maximum poloidal wavenumber of 65.

C Influence of the diagnostic resolution on the frequency spectrum

It was discussed in the main text how the spectral width W_f of the high-k diagnostic is not sensitive to the intrinsic plasma frame turbulence conditions. This begs the question as to what ultimately determines the spectral width in the high-k measurement.

Using the exact same simulation, we perform parametric scans of the synthetic diagnostic for varying values of the wavenumber resolution Δk_Z in figure 18. In the actual experiment, $\Delta k_Z = 2/a_0 = 67$ cm^{-1}, where a_0 is the extent of the microwave beam. In $a)$ is shown how the plasma-frame spectral width value tends to decrease with Δk_Z. This is consistent with our intuition, since lower Δk_Z is implies a smaller wavenumber window contributing to the high-k signal - since nearby modes have similar propagation frequencies, this results in smaller spectral width. The spectral width is shown to be strongly affected by the inclusion of Doppler shift when the wavenumber resolution is large (large wavenumber window), showing W_f can be increased by a factor of ~ 2. However, this effect is negligible for the experimental conditions corresponding to the old high-k system ($\Delta k_z = 67$ cm^{-1}), suggesting the measured spectral width by the old high-k system is similar to the intrinsic plasma frame value.

Acknowledgements

The author would like to thank the whole NSTX team for providing the profile data used for the analysis of the plasma discharge presented here. This work has been supported by US D.O.E. contract DE-AC02-09CH11466. Computer simulations were carried out at the National Energy Research Scientific Computing Center, supported by the Office of Science of the U.S. D.O.E. under Contract No. DE-AC02-05CH11231,

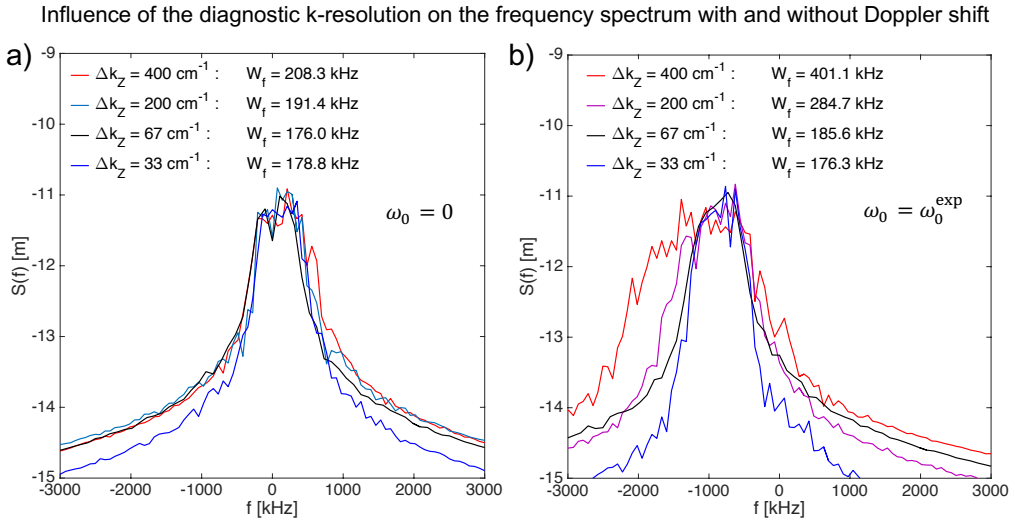

Figure 18: Influence of the diagnostic resolution Δk_Z on the spectral width W_f, in the absence $a)$ and presence of Doppler shift $b)$. These figures show the spectral width W_f can be strongly affected for large values of the diagnostic wavenumber resolution in combination to Doppler shift. However this strong dependence is negligible for the current experimental conditions of $\Delta k_z = 67\ \mathrm{cm}^{-1}$, suggesting the measured spectral width by the old high-k system is similar to the intrinsic plasma frame value.

and at the MIT-PSFC partition of the Engaging cluster at the MGHPCC facility (www.mghpcc.org), which was funded by D.O.E. grant number DE-FG02-91-ER54109. N.F.L. was partially funded by US D.O.E. Grant no. DE-FG02-91ER54109. The digital data for this paper can be found in http://arks.princeton.edu/ark:/88435/dsp018p5

References

[1] W. Guttenfelder, J. L. Peterson, J. Candy, S. M. Kaye, Y. Ren, R. E. Bell, G. W. Hammett, B. P. LeBlanc, D. R. Mikkelsen, W. M. Nevins and H. Yuh, Nucl. Fusion **53**, 093022 (2013).

[2] Y. Ren, E. Belova, N. Gorelenkov, W. Guttenfelder, S. M. Kaye, E. Mazzucato, J. L. Peterson, D. R. Smith, D. Stutman, K. Tritz, W. X. Wang, H. Yuh, C. W. Domier and B. P. LeBlanc, Nucl. Fusion **57**, 072002 (2017).

coupli

[3] T. Brown, J. Menard, P. Titus, A. Zolfaghari, L. El-Guebaly, and L. Mynsberge, "Progess in developping the STFNSF configuration", 25th Symposium on Fusion Engineering (SOFE), San Francisco, June 10-14, 2013.

[4] J.E. Menard, T. Brown, L. El-Guebaly, M. Boyer, J. Canik, B. Colling, R. Raman, Z. Wang, Y. Zhai, P. Buxton, B. Covele, C. D'Angelo, A. Davis, S. Gerhardt, M. Gryaznevich, M. Harb, T.C. Hender, S. Kaye, D. Kingham, M. Kotschenreuther, S. Mahajan, R. Maingi, E. Marriott, E.T. Meier, L. Mynsberge, C. Neumeyer, M. Ono, J.-K. Park, S.A. Sabbagh, V. Soukhanovskii, P. Valanju and R. Woolley, Nucl. Fusion **56**, 106023 (2016).

[5] M. Ono, S. M. Kaye, Y.-K. M. Peng, G. Barnes, W. Blanchard, M. D. Carter, J. Chrzanowski, L. Dudek, R. Ewig, D. Gates, R. E. Hatcher, T. Jarboe, S. C. Jardin, D. Johnson, R. Kaita, M. Kalish, C. E. Kessel, H. W. Kugel, R. Maingi, R. Majeski, J. Manickam, B. McCormack, J. Menard, D. Mueller, B. A. Nelson, B. E. Nelson, C. Neumeyer, G. Oliaro, F. Paoletti, R. Parsells, E. Perry, N. Pomphrey, S. Ramakrishnan, R. Raman, G. Rewoldt, J. Robinson, A. L. Roquemore, P. Ryan, S. Sabbagh, D. Swain, E. J. Synakowski, M. Viola, M. Williams, J. R. Wilson, and the NSTX Team, Nucl. Fusion **40**, 557 (2000).

[6] A. Sykes and the START, NBI, MAST and Theory Teams, Nucl. Fusion **39**, 1271 (1999).

[7] S. M. Kaye, F. M. Levinton, D. Stutman, K. Tritz, H. Yuh, M. G. Bell, R.E. Bell, C. W. Domier, D. Gates, W. Horton, J. Kim, B. P. LeBlanc, N. C. Luhmann Jr, R. Maingi, E. Mazzucato, J. E. Menard, D. Mikkelsen, D. Mueller, H. Park, G. Rewoldt, S. A. Sabbagh, D. R. Smith and W. Wang, Nucl. Fusion **47**, 499 (2007).

[8] S. M. Kaye, S. Gerhardt, W. Guttenfelder, R. Maingi, R. Bell, A. Diallo, B. LeBlanc and M. Podesta, Nucl. Fusion **53**, 063005 (2013).

[9] S. M. Kaye, R. E. Bell, D. Gates, B. P. LeBlanc, F. M. Levinton, J. E. Menard, D. Mueller, G. Rewoldt, S. A. Sabbagh, W. Wang, and H. Yuh, Phys. Rev. Lett **98**, 175002 (2007).

[10] M. Valovič, R. Akers, M. de Bock, J. McCone, L. Garzotti, C. Michael, G. Naylor, A. Patel, C.M. Roach, R. Scannell, M. Turnyanskiy, M. Wisse, W. Guttenfelder, J. Candy and the MAST team, Nucl. Fusion **51**, 073045 (2011).

[11] W. Guttenfelder, J. Candy, Phys. Plasmas **18**, 022506 (2011).

[12] Y. Ren, W. Guttenfelder, S. M. Kaye, E. Mazzucato, R. E. Bell, A. Diallo, C. W. Domier, B. P. LeBlanc, K. C. Lee, D. R. Smith, and H. Yuh, Phys. Plasmas **19**, 056125 (2012).

[13] D. R. Smith, E. Mazzucato, W. Lee, H. K. Park, C. W. Domier, and N. C. Luhmann Jr., Rev. Sci. Instrum., **79**, 123501 (2008).

[14] F. M. Poli, S. Ethier, W. Wang, T. S. Hahm, E. Mazzucato, and D. R. Smith, Phys. Plasmas **17**, 112514 (2010).

[15] W. X. Wang, Z. Lin, W. M. Tang, W. W. Lee, S. Ethier, J. L. V. Lewandowski, G. Rewoldt, T. S. Hahm, and J. Manickam, Phys. Plasmas **13**, 092505 (2006).

[16] J. Ruiz Ruiz, Y. Ren, W. Guttenfelder, A. E. White, S. M. Kaye, B. P. Leblanc, E. Mazzucato, K. C. Lee, C. W. Domier, D. R. Smith, and H. Yuh, Phys. Plasmas **22**, 122501 (2015).

[17] J. Candy and R. E. Waltz, J. Comput. Phys. **186**, (2003) 545.

[18] B.P. LeBlanc, R. E. Bell, D. W. Johnson, D. E. Hoffman, D. C. Long, and R. W. Palladino, Rev. Sci. Instrum, **74**, 1659 (2003).

[19] R. E. Bell, Rev. Sci. Instrum, **74**, 10E902 (2006).

[20] Y. Ren, S. M. Kaye, E. Mazzucato, W. Guttenfelder, R. E. Bell, C.W. Domier, B. P. LeBlanc, K. C. Lee, N. C. Luhmann, Jr., D. R. Smith, and H. Yuh, Phys. Rev. Lett. **106**, 165005 (2011).

[21] R. J. Hawryluk, *Physics of Plasma Close to Thermonuclear Conditions*, (Pergamon, New York, 1981).

[22] E. A. Belli and J. Candy, Plasma Phys. Control. Fusion **50**, 095010 (2008).

[23] R. E. Waltz and R. L. Miller, Phys. Plasmas **6**, 4265 (1999).

[24] N. T. Howard, A. E. White, M. Greenwald, C. Holland, and J. Candy, Phys. Plasmas **21**, 032308 (2014).

[25] N. T. Howard, C. Holland, A. E. White, M. Greenwald, and J. Candy, Phys. Plasmas **21**, 112510 (2014).

[26] N.T. Howard, C. Holland, A.E. White, M. Greenwald and J. Candy, Nucl. Fusion **56**, 014004 (2016).

[27] N. T. Howard, C. Holland, A. E. White, M. Greenwald, J. Candy, and A. J. Creely, Phys. Plasmas **23**, 056109 (2016).

[28] S. Maeyama, Y. Idomura, T.-H. Watanabe, M. Nakata, M. Yagi, N. Miyato, A. Ishizawa, and M. Nunami, Phys. Rev. Lett. **114**, 255002 (2015).

[29] P. W. Terry, M. Greenwald, J.-N. Leboeuf, G. R. McKee, D. R. Mikkelsen, W. M. Nevins, D. E. Newman, D. P. Stotler, Task Group on Verification and Validation, U.S. Burning Plasma Organization, and U.S. Transport Task Force, Phys. Plasmas **15**, 062503 (2008).

[30] M. Greenwald, Physics of Plasmas **17**, 058101 (2010).

[31] E. Mazzucato, Phys. Plasmas **10**, 753 (2003).

[32] E. Mazzucato, Plasma Phys. Control. Fusion **48**, 1749 (2006).

[33] P. Ricci, C. Theiler, A. Fasoli, I. Furno, K. Gustafson, D. Iraji, and J. Loizu, Phys. Plasmas **18**, 032109 (2011).

[34] C. Holland, Phys. Plasmas **23**, 060901 (2016).

[35] C. M. Roach, I. G. Abel, R. J. Akers, W. Arter, M. Barnes, Y. Camenen, F. J. Casson, G. Colyer, J. W. Connor, S. C. Cowley, D. Dickinson, W. Dorland, A. R. Field, W. Guttenfelder, G. W. Hammett, R. J. Hastie, E. Highcock, N. F. Loureiro, A. G. Peeters, M. Reshko, S. Saarelma, A. A. Schekochihin, M. Valovic and H. R. Wilson, Plasma Phys. Control. Fusion **51**, 124020 (2009).

[36] A. M. Dimits G. Bateman, M. A. Beer, B. I. Cohen, W. Dorland, G. W. Hammett, C. Kim, J. E. Kinsey, M. Kotschenreuther, A. H. Kritz, L. L. Lao, J. Mandrekas, W. M. Nevins, S. E. Parker, A. J. Redd, D. E. Shumaker, R. Sydora, and J. Weiland, Physics of Plasmas **7**, 969 (2000).

[37] F. van Wyk, E. G. Highcock, A. A. Schekochihin, C. M. Roach, A. R. Field and W. Dorland, J. Plasma Phys. **82**, 905820609 (2016).

[38] G. M. Staebler, J. Candy, N. T. Howard, and C. Holland, Phys. Plasmas **23**, 062518 (2016).

[39] W. X. Wang, P. H. Diamond, T. S. Hahm, S. Ethier, G. Rewoldt and W. M. Tang, Phys. Plasmas **17**, 072511 (2010).

[40] J. Candy, E. A. Belli, R. V. Bravenec, J. Comp. Phys. **324**, (2016) 73.

[41] M. N. Rosenbluth and F. L. Hinton, Phys. Rev. Lett. **80**, 724 (1998).

[42] F. L. Hinton and M. N. Rosenbluth, Plasma Phys. Control. Fusion **41**, A653 (1999).

[43] Z. Lin, T. S. Hahm, W. W. Lee, W. M. Tang, and P. H. Diamond, Phys. Rev. Lett. **83**, 3645 (1999).

[44] D. R. Mikkelsen and W. Dorland, Phys. Rev. Lett. **101**, 135003 (2008).