Bayesian analysis of turbulent transport coefficients in 2D interchange dominated ExB turbulence involving flow shear

R. Coosemans\textsuperscript{1}, W. Dekeyser, M. Baelmans
Department of Mechanical Engineering, Applied Mechanics and Energy Conversion Section, KU Leuven, Leuven, Belgium
E-mail: reinart.coosemans@kuleuven.be

Abstract. While turbulent transport is known to dominate the radial particle and energy transport in the plasma edge, a self-consistent description of turbulent transport in mean-field transport codes remains lacking. Mean-field closure models based on the turbulent kinetic energy ($k_\perp$) and turbulent enstrophy ($\zeta_\perp$) have recently been proposed to self-consistently model this transport. This contribution analyses the diffusive particle transport relations of these models by means of the Bayesian framework for parameter estimation and model comparison. The reference data includes a set of isothermal simulations that not only include the scrape-off layer (SOL) but also the outer core region (in which a drift wave-like model is activated) and a set of anisothermal SOL simulations, both obtained with the TOKAM2D turbulence code. This analysis shows that the $k_\perp(-\zeta_\perp)$ model does not appropriately capture the diffusion coefficients for these new data sets, presumably due to the strong flows in the diamagnetic direction that appear in these new cases. While flow shear is expected to quench the turbulence and the turbulent transport, its effect was not explicitly taken into account in the earlier $k_\perp(-\zeta_\perp)$ transport models. As flow shear provides a new mechanism for the decorrelation of the turbulence, we propose to introduce an additional time scale $\Omega_s$ in the diffusive transport relation as $D \sim k_\perp/\sqrt{\zeta_\perp + \Omega_s}$. Inspiration is drawn from shear decorrelation times reported in literature to propose several new candidate models, which are then analysed in a Bayesian setting. This allowed identifying irrelevant terms for certain models and to rank all models according to the Bayesian evidence. While the new models accounting for shear do improve the match to the data, significant errors still remain. Also, no single model could be identified that performs best for all data sets.

1. Introduction

In mean-field transport codes, turbulent transport across flux surfaces is commonly approximated by diffusive models with empirically determined diffusion coefficients. To match experiments, these diffusion coefficients typically are modelled as 1D or 2D fields \cite{1-3}, introducing a lot of model flexibility, while not addressing the underlying physics of the turbulence dominating the radial transport. This induces a significant risk of overfitting and limits the reliability of extrapolating to different machines or operating regimes.

Recently, inspiration has been drawn from the Reynolds-Averaged Navier-Stokes approach used in hydrodynamic turbulence modelling \cite{4} to develop mean-field transport models that

\textsuperscript{1} Present address: Celestijnenlaan 300A, 3001 Heverlee, Belgium
start from the underlying characteristics of the turbulence to model the transport coefficients self-consistently [5–7]. In Refs. [5, 6] the turbulent transport is related to the turbulent kinetic energy $k_\perp$, and a model equation for the turbulent kinetic energy is added to the mean-field equation set. Ref. [7] modelled the diffusion coefficient by a dimensionally correct combination of the turbulent kinetic energy and the turbulent enstrophy $\zeta_\perp$ and proposed model equation for both, improving the match with detailed reference data obtained from the TOKAM2D turbulence code [8–10].

Bayesian techniques have recently been applied to estimate unknown transport parameters in mean-field transport modelling in Refs. [11, 12]. There, Bayesian techniques were used to compare the $k_\perp$ model [6] and the $k_\perp - \zeta_\perp$ model in a statistically relevant way. The basic ansatz of these techniques is that a given model does not have a single best or true value for its parameters for a given data set, but rather a probability distribution. The objective of a Bayesian analysis is to infer this distribution. At the same time these distributions can be used to rank the different competing models according to the Bayesian evidence, inherently penalizing models with higher model complexity, hence countering overfitting.

The Bayesian framework developed in Refs. [11, 12] will be applied here to assess the performance of the $k_\perp$ and $k_\perp - \zeta_\perp$ particle transport models under various flow conditions. These models were initially developed for isothermal, 2D interchange dominated scrape-off layer (SOL) ExB drift turbulence. In this contribution on the other hand, anisothermal cases and cases where a core region with drift-wave dynamics has been added will be investigated. It is observed that the earlier models no longer capture the new data to a satisfactory degree. This is assumed to be due to the significant flow in the diamagnetic direction that develops in the new simulations; a phenomenon that was almost absent in the initial isothermal SOL simulations. It is indeed expected that such a flow, and especially a radial flow shear, will reduce the turbulent transport as well as the turbulent kinetic energy [13–17]. Under the hypothesis that it is the novel flow shear phenomenon causing the discrepancy, new models will be proposed that explicitly account for this. This analysis demonstrates the capabilities of Bayesian techniques, and provides guidance to further improve the mean-field turbulent particle transport models.

The remainder of this paper is structured as follows. In section 2, the TOKAM2D turbulence model used to obtain reference data will be introduced and a number of alternative mean-field models for the particle transport closure term will be proposed. Then, the basics of Bayesian analysis will be summarised in section 3. In section 4, these techniques will then be applied to study the performance of both the existing $k_\perp$ and $k_\perp - \zeta_\perp$ particle transport models and the newly suggested ones for the new data. Finally, section 5 will summarise the main findings of this work.

2. Turbulence and mean-field models

This section will first discuss the TOKAM2D code used to obtain the reference data for the different flow cases that are studied. Then, the mean-field transport for these cases will be analysed and mean-field transport models will be proposed.

2.1. 2D $E\times B$ drift turbulence model

To obtain reference data to calibrate the mean field transport models, we use the full anisothermal version of the TOKAM2D code, which describes the turbulence using the following equations [8–10]:

\[ \frac{\partial}{\partial t} \rho + \nabla \cdot (\rho \mathbf{u}) = 0, \]
\[ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\nabla p + \nabla \cdot \mathbf{S} + \mathbf{F}, \]
\[ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \]
\[ \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{v} = -\nabla p + \nabla \cdot \mathbf{T} + \mathbf{F}, \]
In the SOL region, the sources take the form:

$$\frac{\partial n}{\partial t} + V_E \cdot \nabla n = S_n + S_{||, ne} + D_0 \nabla^2 n,$$  

$$\frac{\partial \omega}{\partial t} + V_E \cdot \nabla \omega = -\frac{g p}{n} + \frac{S_{||, ne} - S_{||, ni}}{n} + \nu_0 \nabla^2 \omega,$$  

$$\frac{\partial p_i}{\partial t} + V_E \cdot \nabla p_i = S_{pi} + S_{||, pi} + \chi_0 \nabla^2 p_i + \tau_{ei}^{-1} (p_e - p_i),$$  

$$\frac{\partial p_e}{\partial t} + V_E \cdot \nabla p_e = S_{pe} + S_{||, pe} + \chi_0 \nabla^2 p_e - \tau_{ei}^{-1} (p_e - p_i).$$

In the SOL region, the sources take the form:

$$S_{||, ne} = -\sigma c_s n \exp(\Lambda - \frac{\phi}{T_e}), \quad S_{||, ni} = -\sigma c_s n, \quad S_{||, pi/e} = \frac{2}{3} \gamma_{i/e} T_{i/e} S_{||, ni/e}. \quad (6)$$

while in the core, the sources for the equations are

$$S_{||, ne} = \sigma_{DW} [T_e^{3/2} (\phi - T_e \ln n - T_e) - \langle T_e^{3/2} (\phi - T_e \ln n - T_e) \rangle], \quad (7)$$

$$S_{||, ni} = 0, \quad S_{||, pi/e} = \frac{2}{3} T_{i/e} S_{||, ni/e}. \quad (8)$$

In these equations, $t$ is time, $r$ and $\lambda$ are the radial and diamagnetic directions respectively, $n$ is the density, $p_i$, $p_e$ and $p = p_i + p_e$ the ion, electron and total pressure, $T_i$ and $T_e$ the ion and electron temperature, $\phi$ is the electrostatic potential, $\omega = \nabla^2 \phi$ is the vorticity, $V_E = b \times \nabla \phi$ is the ExB drift velocity where $b$ is the unit vector along the parallel direction. $S_n$, $S_{pi}$, $S_{pe}$ are the particle, ion thermal energy and electron thermal energy sources due to influx from the core region, which have a Gaussian profile in the radial direction and are uniform in the diamagnetic direction and in space. $S_{||, ni/e}$ and $S_{||, pi/e}$ are artificial sources accounting for the parallel dynamics. $c_s = \sqrt{T_i + T_e}$ is the sound speed, $\Lambda$ the sheath potential and $\sigma$, $\sigma_{DW}$ and $g$ are parameters determining the strength of the sheath losses, parallel drift wave (DW) dynamics and magnetic field curvature respectively. $\gamma_i$ and $\gamma_e$ are the sheath heat transmission coefficients for ions and electrons respectively. $\tau_{ei}$ is the collisional ion-electron equilibration time. $\langle x \rangle_\lambda = \int_0^{L_\lambda} x d\lambda / L_\lambda$ denotes the diamagnetic average of $x$, with $L_\lambda$ the diamagnetic length of the domain. Finally, $D_0$ and $\chi_0$ are (small) diffusion coefficients for particles and heat respectively and $\nu_0$ is a viscosity. Note that all quantities in these equations are normalised to the reference ion gyro-period and gyro-radius.

Equations 1-4 are solved on a fine 2D (Cartesian) mesh perpendicular to the magnetic field direction with a small time step by the finite volume version of the TOKAM2D code [9, 10]. Dynamics in the direction parallel to the magnetic field are not explicitly resolved, but are modelled through the artificial sources $S_{||, ni/e}$ and $S_{||, pi/e}$. All the time and length scales of the turbulence are resolved, providing detailed reference data of the turbulence. On radial boundaries, the fluxes of all quantities are forced to zero by Neumann boundary conditions, while periodic boundary conditions are applied on diamagnetic boundaries.

In the core region, the drift wave (DW) terms involving $S_{||, ne}$ have been added to give a 2D representation of DW characteristics. They are derived from the electron momentum balance $\eta_{\parallel} n J_{\parallel} = \nabla \cdot p_e - n \nabla \phi$, where $\eta_\parallel$ is the parallel resistivity. Solving for $J_\parallel$ and taking the divergence, we find

$$\nabla \cdot J_{\parallel} = \nabla \cdot \frac{T_e^{3/2}}{\eta_{\parallel,0}} (T_e \nabla \parallel) \ln n + \nabla \parallel T_e - \nabla \parallel \phi.$$  

(9)
where the resistivity is assumed to scale like $\eta || \sim \eta ||,0 (\frac{T_e}{T_0})^{3/2} \text{ with the reference temperature } T_0 = 1 \text{ (in normalised units).}$ Replacing the parallel gradient operators in this expression with a characteristic parallel wave number $ik||$ and defining $\sigma_{DW} = k_i^2/\eta ||,0$, we obtain $\nabla \cdot J || = \sigma_{DW} T_e^{3/2} (\phi - T_i \ln n - T_e)$. In the final source term, the flux surface (diamagnetic) average of this term is subtracted to ensure zero net divergence of the parallel current on a flux surface. This term leads to the typical DW physics where the potential fluctuations tend to follow the density fluctuations. Note that the $S_{||,ne}$ term used here matches the parallel term in Ref. [17] in the isothermal case and is similar to the parallel term used in the core region in the HESEL code [18].

In the isothermal SOL case studied earlier, only equations 1-2 were solved, while the ion and electron temperature were assumed to be constant in space and time and the DW-like terms were never active. In this scenario, only the interchange instability is present to drive the turbulence. A description of the linear spectrum and the turbulence characteristics for this case can be found in Ref. [10]. Besides this original case, we will also study the isothermal case with a core region where the DW-like terms come into play, and the anisothermal SOL case in which the ion and electron thermal energy equations 3-4 are added such that the dynamics induced by temperature fluctuations are included. For the latter anisothermal SOL case, Refs. [19, 20] showed that an additional instability is present in the model due to the interplay between the electron temperature fluctuations and the parallel source terms modelling the sheath. More information about the influence of these anisothermal dynamics on the turbulence can be found in these references as well. Regarding the isothermal case with core region and SOL, it is clear that the parallel term in the core introduces a DW-like instability in the core region in addition to the interchange instability still at play.

For each of these three flow cases, several simulations were performed with varying input parameters as described in Appendix A. We will refer to the resulting data sets as iso, isoDW, and ani respectively.

2.2. Models for mean-field particle transport

Next, we briefly turn our attention to the mean-field equations corresponding to equations 1-4. The analysis in this paper focuses on the mean-field particle and energy fluxes $\bar{\Gamma}_{E,n}$ and $\bar{\Gamma}_{E,pi/e}$, which can be written as

$$\bar{\Gamma}_{E,n} = \bar{n}V_E = \bar{n}V_E + n'V_E', \quad (10)$$

$$\bar{\Gamma}_{E,pi/e} = \bar{n}T_{i/e}V_E = \bar{T}_{i/e}\bar{\Gamma}_{E,n} + nT_{i/e}'V_E'. \quad (11)$$

Here, both the Reynolds and the Favre decompositions are used [6, 21]. The former decomposes a turbulent quantity $x$ as $x = \bar{x} + x'$ with $\bar{x} = \lim_{T \to \infty} \int_0^T x dt$. The latter uses $x = \bar{x} + x''$ where $\bar{x} = \bar{x}/\bar{n}$ is a Favre or density weighted average. Thus, $\bar{\Gamma}_{E,n}$ is the average ExB particle flux and $\bar{\Gamma}_{E,p}$ the average ExB energy flux. In both expressions, the first term is a mean-field convection term, while the second term is due to turbulent convection. These turbulent convection terms are the main closure terms for which models need to be developed in mean-field transport models. In this contribution we will investigate models for the radial component of these fluxes and consider the component in the diamagnetic direction as a given. Note that this component in the diamagnetic direction was insignificant in the isothermal SOL case considered before, but that it becomes large in the new flow cases. The generation of this flow in the diamagnetic direction is also a topic of active research. Refs. [16, 17] have analysed this in 2D systems for example. However, Refs. [22, 23] found that geodesic curvature effects in 3D toroidal geometries alter the dynamics of zonal flows significantly. Hence, in our analysis, we acknowledge that the shear in the considered 2D system may not be completely representative for realistic 3D
cases, but our working hypothesis is that these 2D systems do allow to study the effect of flow shear on the transport, i.e. that the nature of the transport in the presence of flow shear is not fundamentally changed by the dynamics of the shear generation.

To close the turbulent fluxes, we consider typical diffusion models of the form

\[ n' V'_{E,r} = -D \nabla \bar{n}, \]
\[ n'' T''_{i/e} V''_{E,r} = -\bar{n} \chi_{i/e} \nabla \tilde{T}_{i/e}. \]  \hfill (12)

Figure 1 shows the radial profiles of the turbulent transport coefficients for a representative simulation of each of the three studied flow cases. The left plot in this figure shows that the diffusion coefficient for a typical iso simulation has a very flat profile. In the isoDW case, shown in the middle plot, the profile in the SOL changed little, but there is a large suppression of the diffusion coefficient in the core region. The right plot shows that the diffusion coefficient develops a clear radial profile in the ani case. This plot also shows that the turbulent heat transfer coefficients seem to be proportional to the particle diffusion coefficient with proportionality factors close to unity, i.e. \( \chi_i \sim \chi_e \sim D \). Since this is also observed in the other ani simulations, we will focus on models for the particle diffusion coefficient \( D \) only, assuming that this will readily yield a model for \( \chi_i \) and \( \chi_e \) as well.

Figure 1. Radial profiles of the turbulent transport coefficients for an iso simulation (left), an isoDW simulation (middle) and an ani simulation (right). The left and right plots consider the SOL only, while the separatrix lies at \( r = 0 \) in the middle plot.

Ref. [6] studied this particle diffusion coefficient in the isothermal SOL case only and proposed the \( k_\perp \) model:

\[ D = C_D k_\perp, \]  \hfill (13)

where the turbulent kinetic energy is defined as \( k_\perp = \langle n V_{E,r}^2 \rangle / \bar{n} \) with \( C_D \) a free parameter. A value of \( C_D = 23.9 \) was found by means of a regression analysis for a set of TOKAM2D simulations. Still considering only the SOL under isothermal conditions, Ref. [7] proposed the \( k_\perp - \zeta_\perp \) model as a refinement of the previous model:

\[ D = C_D \frac{k_\perp}{\sqrt{\zeta_\perp}}, \]  \hfill (14)

where the turbulent enstrophy is defined as \( \bar{n} \zeta_\perp = \langle n \omega_{i/e}^2 \rangle / 2 \). A regression analysis for a (slightly different) set of TOKAM2D data yielded \( C_D = 7.6 \).

A significant effect coming into play that was almost absent in this iso flow case is flow shear. In the new flow cases, this is especially pronounced in the core region of the isoDW simulations. Literature indicates that mean-flow shear changes the characteristics of the turbulence and decreases the transport [13–17]. It is expected that both the intensity of the turbulence, thus
$k_{\perp}$ and $\zeta_{\perp}$, and the magnitude of the transport for a given turbulence intensity, thus $D$ at fixed $k_{\perp}$ and $\zeta_{\perp}$ decreases. In this contribution we will only investigate the latter effect.

It might be hoped that the effect of flow shear would inherently be captured by the scales used in the $k_{\perp} - \zeta_{\perp}$ model, however, it will be shown that this model no longer performs satisfactorily for the new flow cases. The ansatz followed in this contribution is that this reduced accuracy is due to the presence of radial shear of the diamagnetic component of the ExB velocity, $S_{\lambda} = \partial_r V_{E,\lambda}$.

A first modification of models (13) and (14) that will be tested is to replace the total perpendicular turbulent kinetic energy $k_{\perp}$ by the radial turbulent kinetic energy $k_r = nV_{E,\perp}^2/n$. The idea behind this modified version is that the turbulent kinetic energy in the radial velocity component may be more strongly related to the average radial transport than the total perpendicular kinetic energy, especially in cases where there is strong shear, presumably affecting the turbulent kinetic energy in the diamagnetic velocity component $k_{\lambda} = nV_{E,\perp}^2/n$.

Another adjustment proposed here is to take the effect of the flow shear explicitly into account, by adding a frequency scale accounting for the flow shear $\Omega_S$ to the regular $k_{\perp} - \zeta_{\perp}$ model:

$$D = C_D \frac{k_{\perp}}{\sqrt{\zeta_{\perp} + C_S \Omega_S}},$$

with $C_S$ a proportionality constant to be determined. The rationale is that $\sqrt{\zeta_{\perp}}$ provides a characteristic frequency for the temporal decorrelation of eddies (eddy turn-over frequency), while shear provides an additional decorrelation mechanism. As shear becomes large, the latter starts dominating the former and decreases the turbulent diffusion coefficient for a given turbulence intensity as indicated by $k_{\perp}$ and $\zeta_{\perp}$. This newly proposed model is in accordance with Refs. [13, 14] which indicate that fluctuation suppression occurs when the shearing rate becomes larger than the ambient turbulence decorrelation rate.

Shear decorrelation rates indicated in literature make a natural choice for the shear frequency scale $\Omega_S$. A very basic frequency scale is $\Omega_S \sim |S_{\lambda}|$. A slightly more complicated decorrelation rate is based on the time it takes for two fluid elements a distance $\Delta r$ away to be decorrelated by moving a diamagnetic wavelength $K_{\lambda}$ away in the diamagnetic direction, resulting in $\Omega_S \sim K_{\lambda} |S_{\lambda}| / \Delta r$ [13, 15]. Finally, the interplay between regular radial decorrelation and shear flow decorrelation may result in a hybrid decorrelation time $\Omega_S \sim (D K_{\lambda}^2 S_{\lambda}^2)^{1/3}$ [13, 15].

These scalings only strictly apply for transport of passive scalars in a constant shear flow. However, here we use them to model the effect of flow shear which may vary in time on the density transport. To this end, we first define the mean-flow shear $S_{\lambda,\text{mean}} = \partial_r \bar{V}_{E,\lambda}$, which is the shear based on the Favre average velocity in the diamagnetic direction. In order to convert the first shear time scale to a mean-field one we simply take $\Omega_S \sim |S_{\lambda,\text{mean}}|$. This will be referred to as the simple shear model in the following. For the second time scale we have to use an approximation for $K_{\lambda}$ and a relevant scale for $\Delta r$. We construct length scales for both based on the radial and diamagnetic turbulent kinetic energy, which nicely fits in the modelling framework constructed in Refs. [6, 7]. This then yields $\Omega_S \sim |S_{\lambda,\text{mean}}| \sqrt{k_r/k_{\lambda}}$, which will be referred to as the strong shear model. For the hybrid decorrelation time model, a time scale is needed in addition to the $k_{\lambda}$ velocity scale to construct an approximation for $K_{\lambda}$. If the turbulent enstrophy is used for that, we find $K_{\lambda} \sim \sqrt{\zeta_{\perp}/k_{\lambda}}$ and then $\Omega_S \sim (D S_{\lambda,\text{mean}}^2 \zeta_{\perp}/k_{\lambda})^{1/3}$. Note that the diffusion coefficient is now ultimately present in the model for $D$ itself. In a mean-field code that is solved iteratively, this should not pose a problem. In case the shear frequency is dominant ($D \sim k/\Omega_S$) the diffusion coefficient can be eliminated to find $D \sim k^{3/4}/\sqrt{|S_{\lambda,\text{mean}}| K_{\lambda}}$ as was done in Ref. [15]. This then yields $\Omega_S \sim \sqrt{|S_{\lambda,\text{mean}}| (k_{\lambda}/k_{\lambda})^{1/4}}$ for the resulting mean-field time scale. Note that in the previous derivation for the hybrid decorrelation frequency, the “regular” turbulent enstrophy time scale of the turbulence was used. It may however be more consistent to use
the shear decorrelation frequency instead. Doing so, we find \( \Omega_S \simDS^2/k_\lambda \). When this is then in turn filled out in the diffusion relation for the case of dominant shear frequency, we finally obtain \( D \sim \sqrt{k_\lambda S_{\lambda,\text{mean}}^2} \). This can be seen to be equivalent to the regime where shear is dominant for the strong shear model (when \( k = k_r \)). Hence, the hybrid decorrelation time reduces to the strong shear decorrelation time in the strong shear limit in these models. The diffusion models resulting from these four mean-flow shear time scales are summarised in table 1.

3. Bayesian analysis methodology

The mean-field transport model development is significantly complicated by the additional dynamics that come into play in the two new flow cases which are investigated in this paper. In order to deal with this increased complexity, we apply Bayesian techniques for parameter estimation and model comparison. The solid statistical basis of this approach is expected to give a clearer view of which quantities are relevant and which models perform better compared to the least squares regression techniques used in earlier studies [6, 7].

3.1. Basics of Bayesian theory

The objective of a Bayesian analysis is to infer about the probability distribution of the model parameters, which provides much more information than a single parameter value resulting from a classical least squares regression. In addition, this methodology also allows to rank different competing models according to the Bayesian evidence.

The basic identity underlying Bayesian techniques is Bayes’ rule:

\[
P(\theta|D, M) = \frac{\pi(\theta|M) L(D|\theta, M)}{L(D|M)},
\]

where the posterior \( P(\theta|D, M) \) is the probability of the parameters \( \theta \) after observing the data \( D \) for a certain model \( M \), the prior \( \pi(\theta|M) \) is the probability of the parameters before observing the data for the model, the likelihood \( L(D|\theta, M) \) indicates how likely it is to observe the data given a specific value of the parameters and the model, and the Bayesian evidence is the marginal likelihood of observing the data given the model no matter the parameters [11, 12, 24–26].

The main objective of a Bayesian analysis for parameter estimation is to infer the posterior distribution of the parameters \( P(\theta|D, M) \) from the data. Secondly, the model evidence \( L(D|M) \) can be evaluated for model comparison. In order to achieve this, a likelihood function and a prior need to be defined. The former is basically a model for the probability distribution describing the error between the data and the model. If we consider a model \( M : f(x, \theta) \) (which is a function of the input data \( x \)) that approximates the (output) data \( D : y \), the error is defined as \( \epsilon = y - f \). This error is typically assumed to follow a normal distribution with zero mean \( \epsilon \sim N(0, \sigma) \). The likelihood function is then this distribution expressed as a function of the parameters, yielding for a single data point:

\[
L(D_i|\theta, M) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left(-\frac{(y_i - f(x_i, \theta))^2}{2\sigma_i^2}\right).
\]

For a data set consisting of multiple points and allowing for possible correlations between the data, we get

\[
L(D|\theta, M) = \frac{1}{(2\pi)^{n/2} \det \Sigma^{1/2}} \exp\left(-\frac{1}{2} \epsilon^T \Sigma^{-1} \epsilon\right),
\]

where it is understood that \( \epsilon \) is a vector grouping the errors \( \epsilon_i \) of all individual data points and \( \Sigma \) is the covariance matrix of the error [11, 12, 26]. Hence, in order to evaluate the likelihood
function for a single value of the parameters $\theta$, a forward model evaluation of $f(x, \theta)$ is required for all data points. Note that while the likelihood function is a proper probability density as a function of the model error, it is not a probability density when viewed as a function of the parameters.

The prior $\pi(\theta|M)$ then describes what is known about the parameters prior to observing the data. As the models developed in this work are novel such that no prior information is available, it is opted to use wide Gaussian prior distributions in order not to exclude any values of the parameters from the start.

As is evident from equation (16), the product of the likelihood and the prior $L(D|\theta,M)\pi(\theta|M)$ determines the shape of the posterior distribution $P(\theta|D,M)$, which is the end goal for the parameter estimation since the evidence is a scalar which can be considered to be a normalisation factor in equation (16). This posterior provides a comprehensive description of the information the data provides on the parameters: the regions in parameter space that are most likely, how large the spread is on the parameters and the correlation between the different parameters of the model $M$.

The evidence $L(D|M)$ is of particular interest to compare different competing models attempting to explain the data though. Defined as

$$L(D|M) = \int \pi(\theta|M)L(D|\theta,M)d\theta, \tag{19}$$

it is independent of the value of the parameters. Note that the evidence is not a probability, since the likelihood is not a probability distribution in terms of the parameters. However, the Bayesian evidence is directly related to the probability of a model given the data:

$$P(M|D) = \frac{P(M)L(D|M)}{L(D)}, \tag{20}$$

where $P(M)$ is the prior probability for the model before observing the data (independent of the parameters) and $L(D)$ is a normalisation factor that purely depends on the data. Hence, if there is no prior information $P(M)$ about which model is more probable, the ratio of the evidence of two models $M_1$ and $M_2$ is equal to the relative probability of both models given the data [24, 25]:

$$\frac{P(M_1|D)}{P(M_2|D)} = \frac{L(D|M_1)}{L(D|M_2)}. \tag{21}$$

Thus, while the evidence for a single model does not have a clear meaning (as it is not a probability), the ratio of the evidence between competing models directly indicates which model is to be preferred. Note that this only holds for models trying to explain the same data, the evidence cannot be used to compare models that explain different data sets.

To highlight the main features of Bayesian theory, we briefly compare to a classical least square regression approach. If the model error $\epsilon$ is assumed to have a Gaussian distribution as in equation (18) and when the covariance matrix would only have diagonal elements that are uniform ($\Sigma = \Sigma_{ii} = \sigma^2$), it can be seen that maximising this likelihood is identical to a classical least squares regression. So, a first generalisation is that the likelihood function also allows more complex distributions for the error if desired. In a second step, the likelihood function is multiplied with the prior distribution such that any available prior information could be self-consistently incorporated. Maximising this product yields the maximum a posteriori (MAP) estimate for the parameters. In addition, the distribution of the parameters formed by the product $L(D|\theta,M)\pi(\theta|M)$ is proportional to the probability density function for the parameters, hence including much more information than the MAP point-estimate itself. Comparing different
models based on the value of $\mathcal{L}(\mathcal{D}|\theta, \mathcal{M})\pi(\theta|\mathcal{M})$ at MAP would still only compare models for a single parameter setting, as the least square comparison does. Instead, the Bayesian evidence $\mathcal{L}(\mathcal{D}|\mathcal{M})$ integrates this product of likelihood and prior over the parameter space of the models in order to effectively compare the models for any setting of the parameters allowed by the prior. This “prior-weighed” integration of the likelihood function over parameter space adds what is called an “Occam factor”, penalizing more complex models to the basic (least squares-like) goodness of fit statistic included in the likelihood function [24, 25].

3.2. Numerical techniques for Bayesian inference

In order to infer about the complex posterior distribution $\mathcal{P}(\theta|\mathcal{D}, \mathcal{M})$, Markov Chain Monte Carlo (MCMC) techniques are used in this work. The idea is to generate a chain of samples which converges to a statistical steady state, after which the samples are distributed as the target distribution. More in particular, the Adaptive Metropolis-Hastings (AMH) algorithm [11, 12, 27] is used. A typical step of the algorithm for sampling from a generic target distribution $h(\theta)$ is schematically represented below:

- Given $\theta_i$, generate $\theta'$ according to the proposal distribution $q(\theta'|\theta_i)$
- Calculate the acceptance probability $\alpha = \frac{h(\theta')q(\theta'|\theta_i)}{h(\theta_i)q(\theta_i|\theta')}$
- generate $u \sim U(0,1)$
  - if $\alpha > u$ accept sample, $\theta_{i+1} = \theta'$
  - if $\alpha < u$ reject sample, $\theta_{i+1} = \theta_i$

In this algorithm, $q(\theta'|\theta_i)$ is the proposal distribution describing the probability of where the next proposal parameter will be located given the previous accepted sample. The choice of this distribution is crucial to the efficiency of the algorithm. If this distribution is very narrow, samples will easily be accepted as they are likely to be close to the previous accepted sample, however, it will take a long time for the chain to move to other regions and explore the whole target distribution. If the proposal distribution is very wide on the other hand, the accepted samples will quickly cover the entire target distribution, but a high number proposal samples will be rejected before one is accepted since they are likely to be far off in low probability regions. In order to increase the efficiency of the algorithm, the proposal distribution is adjusted during runtime in the AMH algorithm. The previously accepted samples are used to estimate the covariance matrix of the target distribution, which is then optimally scaled and used as the covariance matrix of the Gaussian proposal function centered around the last accepted sample [11, 12, 27, 28]. This algorithm is used for Bayesian inference by sampling for $h(\theta) \sim \pi(\theta|\mathcal{M})\mathcal{L}(\mathcal{D}|\theta, \mathcal{M})$.

Chib’s method, described in Ref. [29], is then used to calculate the Bayesian evidence from the samples the AMH algorithm has generated (as the evidence is not directly an output of the latter). More information about the specifics of the algorithms used in this paper can be found in Refs.[11, 12].

4. Transport coefficient modelling using Bayesian analysis

In this section, the Bayesian framework for parameter estimation and model comparison used to study the particle transport models in the new isoDW and ani flow cases and compare to results of the original iso cases. First, the $k_\perp$ and $k_\perp - \zeta_\perp$ models developed in Refs. [6, 7] will be assessed in section 4.1, and then the new models explicitly taking flow shear into account will analysed in section 4.2.

4.1. Assessment of $k_\perp$ and $k_\perp - \zeta_\perp$ models

As data for the Bayesian analysis, we consider the 1D profiles of the TOKAM2D reference data obtained by averaging in time (from the onset of the statistical steady state) and in the
diamagnetic direction (which is a symmetry direction). Hence, the output data $y$ are the radial profiles of the diffusion coefficient $D$, and the input data $x$ are the radial profiles of turbulent kinetic energy $k_\perp$ and the turbulent enstrophy $\zeta_\perp$.

Describing the model error $\epsilon$ in order to construct the likelihood function is a delicate matter. This error is not only composed of the statistical error that remains on the averaged output data $y$ but also includes the error on the model $f(x,\theta)$. The latter is composed of an error because the model form $f(x,\theta)$ is not perfect (model inadequacy), an error caused by propagating the statistical error on the averaged input data $x$ through the model and a numerical error in solving the model (which should however be negligible for the algebraic models considered here) [30]. Since very little is known about most of these error contributions, we chose to aggregate them all and assume the error at every data point has a Gaussian distribution, which is consistent with a minimum entropy principle [26, 31]. The standard deviation of the error is assumed to scale with the magnitude of the data such that $\epsilon_i \sim N(0, \sigma y_i)$, where $\sigma$ is assumed to be the same for all data points. As this “relative standard deviation” $\sigma$ is not known a priori, it is an additional parameter that will be inferred. The total likelihood function is then the product of the likelihood functions of all individual data points. For the priors of $C_D$ and $\sigma$ we chose wide distributions in order not to exclude any parameter values a priori:

$$
\pi(C_D|M) \sim 2N(0, 100) \quad \text{for} \quad C_D > 0,
$$

$$
\pi(\sigma|M) \sim \frac{1}{\sigma} \quad \text{for} \quad \sigma > 0.
$$

The results of the Bayesian analysis for this set-up are summarised in the first rows of table 1. The maximum a posteriori (MAP) value of the parameters of both models as well as the logarithm of their evidence is shown. Both the original models and the variant in which the total perpendicular turbulent kinetic energy $k_\perp$ is replaced by the radial turbulent kinetic energy $k_r$ are studied. Results are shown for the three data sets described in Appendix A. Looking at the MAP values of the $C_D$ coefficient for the regular $k_\perp$ and $k_\perp - \zeta_\perp$ models in the iso case, it can be seen that these are rather close to the values of respectively 23.9 and 7.6 that were found by means of regression analysis in Refs. [6, 7]. Considering the logevidence of the models, it is clear that for all data sets the $k_\perp - \zeta_\perp$ model outperforms the $k_\perp$ model, as was also found in Refs. [11, 12] and qualitatively in Ref. [7]. In addition, the models where $k_r$ is used instead of $k_\perp$ perform better in the majority of cases. Remember that the evidence also depends on the data, such that it can only be meaningfully compared for models which have the same output data $y$. Hence only models within (double) columns under iso, isoDW and ani that use the same data set can be compared with each other. Also, since the (natural) logarithm of the evidence is reported, a difference of about 5 between two models can already be considered to be decisive in selecting one model over another [11, 12, 24]. Now looking at the estimate of the relative standard deviation $\sigma$ of the error, it is clear that this is significantly smaller for the iso data than for the ani data, and much smaller than for the isoDW data. This indicates that the models perform much less well for the new data sets.

Figures 2 and 3 support the observation that the $k_\perp - \zeta_\perp$ model is unable to capture the trends in the new data sets. Figure 2 compares the radial profile for the diffusion coefficient for representative simulations to the models proposed here. It can be seen that the $k_\perp - \zeta_\perp$ model captures the profiles very well for the iso and ani (SOL) data sets, but that it fails to properly predict the diffusion coefficient in the isoDW dataset. The model seems to be making a trade-off between having a high enough diffusion coefficient in the SOL and capturing the suppressed value in the core region. Figure 3 then gives an overview of the model’s performance in parameter space. It shows a scatter plot of the model value for $D$ on the vertical axis versus the observed TOKAM2D value on the horizontal axis. Each circle in the figure represents a single TOKAM2D simulation (for a certain parameter set $g$, $\sigma$, $\nu$, ...), which is now also radially
Table 1. Bayesian analysis of the proposed models for the particle diffusion coefficients for the iso, isoDW and ani datasets. Parameter values are reported at MAP.

| Model for $D$ | Model for $D$ | parameter | $k_\perp$ | $k_r$ | $k_\perp$ | $k_r$ | $k_\perp$ | $k_r$ |
|---------------|---------------|-----------|----------|--------|----------|--------|----------|--------|
| $C_D \sqrt{k}$ | $C_D \sqrt{k}$ | $C_D[-]$ | 26.6     | 32.0   | 25.8     | 31.5   | 15.8     | 19.7   |
|               | $\sigma[\%]$ | 22.3     | 20.3     | 74.6    | 65.8    | 29.0    | 29.2    |
|               | logevidence   | 403      | 800      | -2208   | -1673   | 1349    | 1324    |
| $C_D \frac{k}{\sqrt{\zeta_\perp}}$ | $C_D \frac{k}{\sqrt{\zeta_\perp}}$ | $C_D[-]$ | 7.82     | 11.4   | 5.83     | 8.81   | 4.73     | 7.32   |
|               | $\sigma[\%]$ | 9.33     | 7.74     | 67.8    | 57.6    | 26.7    | 24.1    |
|               | logevidence   | 3974     | 4731     | -2104   | -1319   | 1642    | 2051    |
| $\Omega_S = C_S|S_{\lambda,mean}|$ | $\Omega_S = C_S|S_{\lambda,mean}|$ | $C_D[-]$ | 7.90     | 11.4   | 8.04     | 10.8   | 5.74     | 8.55   |
|               | $C_S[-]$      | 2.40     | 0.938    | 2.79    | 1.12    | 19.3    | 15.0    |
|               | $\sigma[\%]$ | 9.30     | 7.74     | 48.3    | 47.7    | 24.0    | 22.0    |
|               | logevidence   | 3983     | 4729     | -582    | -546    | 2079    | 2406    |
| $\Omega_S = C_S|S_{\lambda,mean}|\sqrt{\frac{k_r}{k_\perp}}$ | $\Omega_S = C_S|S_{\lambda,mean}|\sqrt{\frac{k_r}{k_\perp}}$ | $C_D[-]$ | 7.85     | 11.4   | 8.03     | 11.2   | 5.67     | 8.55   |
|               | $C_S[-]$      | 0.635    | 0.392    | 2.75    | 1.43    | 13.4    | 11.2    |
|               | $\sigma[\%]$ | 9.33     | 7.74     | 52.2    | 47.1    | 24.5    | 22.2    |
|               | logevidence   | 3971     | 4726     | -836    | -491    | 1995    | 2376    |
| $\Omega_S = C_S\left(D_{\alpha,mean}\zeta_\perp\right)^{1/3}$ | $\Omega_S = C_S\left(D_{\alpha,mean}\zeta_\perp\right)^{1/3}$ | $C_D[-]$ | 7.82     | 11.4   | 8.28     | 10.9   | 5.49     | 8.27   |
|               | $C_S[-]$      | 1.18E-3  | 9.75E-5  | 1.07    | 0.507   | 1.35    | 1.10    |
|               | $\sigma[\%]$ | 9.33     | 7.74     | 60.3    | 53.2    | 26.1    | 23.5    |
|               | logevidence   | 3966     | 4722     | -1429   | -947    | 1748    | 2143    |
| $\Omega_S = C_S\left(S_{\alpha,mean}\zeta_\perp\right)^{1/4}$ | $\Omega_S = C_S\left(S_{\alpha,mean}\zeta_\perp\right)^{1/4}$ | $C_D[-]$ | 7.92     | 11.4   | 10.78    | 13.32  | 6.55     | 9.57   |
|               | $C_S[-]$      | 0.154    | 8.16E-2  | 3.02    | 1.67    | 3.03    | 2.70    |
|               | $\sigma[\%]$ | 9.32     | 7.74     | 52.8    | 47.8    | 24.7    | 22.4    |
|               | logevidence   | 3973     | 4726     | -869    | -527    | 1968    | 2330    |
| $\Omega_S = C_S\left(S_{\lambda,mean}\zeta_\perp\right)^{1/4}$ | $\Omega_S = C_S\left(S_{\lambda,mean}\zeta_\perp\right)^{1/4}$ | $C_D[-]$ | 7.92     | 11.4   | 9.01     | 11.7   | 5.68     | 8.55   |
|               | $C_S[-]$      | 4.52E-3  | 2.91E-3  | 1.93    | 1.13    | 13.38   | 11.25   |
|               | $\sigma[\%]$ | 9.34     | 7.75     | 52.0    | 47.2    | 24.5    | 22.2    |
|               | logevidence   | 3967     | 4720     | -820    | -491    | 1989    | 2369    |

averaged. While this radial averaging is not a part of the model development since the profiles are non-uniform, this kind of figures does give an overview of the performance in parameter space. Note that the classical coefficient of determination is also reported in these figures. It is defined as $R^2 = 1 - \sum (f_i - y_i)^2 / \sum (y_i - y_m)^2$, where $y_m = \sum y_i$. The figure shows that $k_\perp - \zeta_\perp$ performs very well also in parameter space for the iso dataset, while quite some scatter remains on the ani data set and that it performs much less well for the isoDW data.

In order to further analyse the performance of the $k_\perp - \zeta_\perp$ models in the isoDW flow case and to investigate the significance of the shear models for this case later on, we also compare with “2 zone” models shown in the third row of table 1. The idea for these models is that the dynamics in the core and those in the SOL may be inherently different because of the different parallel dynamics. As such, it could be imagined that both regions actually follow the diffusion relation of the $k_\perp - \zeta_\perp$ model, but that different model constants should be allowed in the core.
Figure 2. Radial profiles of the particle diffusion coefficient for an iso simulation (left), an isoDW simulation (middle) and an ani simulation (right). TOKAM2D diffusion coefficient shown in blue, other lines show different models. The left and right plots consider the SOL only, while the separatrix lies at $r = 0$ in the middle plot.

Figure 3. Scatter plot of the particle diffusion coefficient for iso (left), isoDW (middle) and ani (right) data sets for different models.

...and the SOL. Table 1 shows the evidence of this model to be much higher. The middle plot of figure 2 clearly shows that this model now captures the behaviour in the SOL very well, while the fit in the core region is also improved. The Figure 3 indicates that the 2 zone model performs much better in parameter space as well.

It can also seen that the $C_D$ parameter in the SOL of this model is close to that of the iso flow case as would be expected for a well performing model. However, the $C_D$ coefficient in the core region and that of the ani case is remarkably lower than that in the other cases. We interpret this as an indication that the turbulence characteristics and related transport differ between the flow cases, and an additional confirmation that the simple $k_\perp(\perp\zeta)$ models are insufficient to describe all flow cases.

4.2. Assessment of shear-adjusted particle transport models

The next five rows in table 1 show the shear-adjusted models suggested in section 2.2 and the Bayesian analysis results for them. The same set-up and priors have been used as for the analysis of the $k_\perp$ and $k_\perp - \zeta_\perp$ models. In addition, a new prior needs to be specified for the $C_S$ parameters, which is likewise chosen to be uninformative as $\pi(C_S|M) \sim 2N(0, 100)$ for $C_S > 0$.

The table shows that the $k_r - \zeta_\perp$ model has the highest evidence for the iso data set. Because this model works very well already, adding a new term to the model to incorporate flow shear does not bring a significantly better fit while increasing its prior parameter volume. As a result, adding a shear time scale tends to decrease the model evidence. Note also that the Bayesian techniques automatically predict very low values of the new, insignificant term for most models as expected.

Looking at the ani results next, it is observed that here the shear models do appreciably...
improve the evidence and reduce the model error. Surprisingly, it is the simplest shear model with $k_r$ in the denominator that works best in this case. It is also interesting to notice that the bottom model, which includes a combination of the strong shear decorrelation timescale and the hybrid decorrelation timescale to account for the possible influence of both, does not bring an improvement with respect to the model with the strong shear model time scale only. Note that the model parameters tend to the parameters of the model with the higher evidence of the both without making a compromise. In addition, the model evidence of this combined model is lower than that of the strong shear case because of its higher complexity and number of fitting parameters.

Finally, including the flow shear in the models for the $isoDW$ data significantly increases the evidence and decreases the model error over the regular $k_{\perp} - \zeta_{\perp}$ model as expected. Interestingly, these shear models also provide an improvement over the 2 zone models, indicating that they are more significant. Here, the strong shear model with $k_r$ in the denominator has the highest evidence together with the model with two timescales. In the latter, it seems the model does combine both time scales to improve its results, which presumably has more impact for this case since the remaining error is larger. In general, it has to be remarked that the error on the models for this $isoDW$ data remains large.

Figures 2 and 3 show the performance of some of the better models in terms of radial profiles and in parameter space. The left and right plots of figure 2 show that all models manage to capture the radial profiles very well for the $iso$ and $ani$ data sets. The middle figure shows that the strong shear model and the 2 zone model manage to capture the base level of the suppressed diffusion coefficient in the core region rather well, but that both models show significant peaks that are not present in the data. Figure 3 then shows that all models correctly identify the trend in parameter space for the isothermal data as expected. The shear models manage to improve the fit for the anisothermal dataset as well, although a large improvement is already achieved by using $k_r$ instead of $k_{\perp}$ in the denominator. Much more scatter remains than on the $iso$ data though. The middle scatter plot in figure 3 seems to indicate that both the shear model and the 2 zone model manage to capture the trends in parameter space rather well, with the 2 zone model performing better according to the classical $R^2$ value. It has to be remarked though that this plot pertains to the radial average of the entire domain, both outer core region and SOL. If a similar scatter plot is made for the core region only, the models perform much less good and the $R^2$ value is seriously reduced. Hence, the $R^2$ value and the Bayesian evidence yield largely the same ranking for the models over the different data sets, but for the $isoDW$ flow case, the ranking of the strong shear model and the 2 zone model is reversed. Lastly, it is worth noting that the results of the $isoDW$ shear models seem to be applicable to the $iso$ data as well. This is indicated by the last set of data in the left plot of figures 2 and 3, which represent the strong shear model evaluated with the MAP value of the parameters found for the $isoDW$ data. This seems to be compatible with the $iso$ results as well, both on the radial profiles and across parameter space. Note that we did not include results from the simple shear model in any of these plots to facilitate interpretation. The simple shear model results are very similar those of the strong shear model in most cases, but they are slightly less good for the $isoDW$ data in figure 3.

An important consideration and possible (partial) explanation for the mediocre performance of the shear-adjusted diffusion models reported here may be in the time dependent nature of the flow shear. Assume for example that an exact expression for the relation between instantaneous transport and instantaneous flow shear would be available, $D(t) = f(S_{\lambda,mean}(t))$. If this relation is not linear, $\bar{D} \neq f(\bar{S}_{\lambda})$. In our models, we use the mean flow shear as a proxy for the shear, which may impact the performance of the mean field models if fluctuations are large.

Finally we remark that the influence of the Bayesian parameters and the statistical framework should be kept in mind. The results of the Bayesian analysis do not only depend on the physical
model (the diffusion models that are proposed), but also on the statistical model: the form of the likelihood function that is chosen, the correlation length and the prior. The width of the prior in particular determines how strongly more complex models are penalised. Inferences with various correlation lengths showed that the value of the parameters depends on it in a rather irregular way, especially for those models that seem to perform less good. As such, we chose to use the results where no correlation length is used. The value of the correlation length did not seem to change the ranking of the different models according to their evidence though. In this context, it may be envisaged to smooth the results in the core region of the isoDW data since the peaks are probably due to the rather small scale specifics of the flow field which are hard to capture in a mean-field model. However, care should be taken for the interaction with the statistical model of the Bayesian framework.

5. Conclusion
In this contribution, mean-field models for the particle diffusion coefficient are analysed based on detailed reference data from the TOKAM2D turbulence code. Isothermal SOL cases, isothermal simulations with SOL and outer core region and anisothermal SOL simulations are considered. This model development process is supported by a Bayesian framework for parameter estimation and model comparison. These Bayesian techniques allowed estimating the combined posterior distribution of the proposed models, identifying terms which are irrelevant in the models and ranking the different models according to their evidence. This Bayesian evidence is especially relevant since it automatically penalises more complex models hence guarding against over fitting.

This analysis has shown that the existing \( k_{\perp} \) and \( k_{\perp} - \zeta_{\perp} \) particle transport models proposed in Refs. [6, 7] do not capture the trends in the new simulations with a core region and in the anisothermal simulation. Flow shear is assumed to be the origin of this deficiency and the \( k_{\perp} - \zeta_{\perp} \) diffusion model is adapted by incorporating shear decorrelation frequencies inspired by literature to explicitly take this effect into account.

The shear-corrected models provide a large enhancement over the earlier models, but it was not possible to identify a single model as being the best for all data. Using the radial turbulent kinetic energy instead of the total perpendicular kinetic energy provided a significant improvement of the results for all data sets, be it in the original or in the shear corrected models. However, considerable room for improvement certainly does remain for the anisothermal simulations and the core region.

Based on this analysis, it is plausible that the discrepancies of the earlier models are indeed due to flow shear, but it cannot be ruled out that other effects are at play instead or in addition. The model parameters in the anisothermal case are systematically lower than those of the isothermal cases for example. This may be due to the inherent differences between the regimes, but might also indicate that certain anisothermal effects are still missing. This will be analysed in more detail in future work. Next, forward simulations will have to prove if the models developed here effectively captured causal relations, or rather correlations without their common origin. Further research will also focus on investigating the balances of the turbulent kinetic energy and turbulent enstrophy for the new flow cases following the methodology demonstrated in Refs. [6, 7]. Another research track is to further extend the physical model of the TOKAM2D simulations that are studied. Up to now, the diamagnetic drift contribution to the vorticity has been neglected, while literature indicates this may significantly influence the dynamics [20]. Also, anisothermal cases including a core region and a SOL could be investigated.

Acknowledgments
R. Coosemans is funded by a strategic basic research grant of the Research Foundation Flanders (FWO), file number 48697.
This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

The computational resources and services used in this work were provided by the VSC (Flemish Supercomputer Center), funded by the Research Foundation Flanders (FWO) and the Flemish Government – department EWI.

The authors would like to thank R. De Wolf for providing the Bayesian framework implemented in Matlab and for the support in using it.

Appendix A. TOKAM2D simulation parameters
This appendix summarises the TOKAM2D simulation parameters used in the three datasets used in this paper. In all simulations, the grid size was chosen to be $\Delta x = \Delta y = 1$ gyro radius and fringe regions are applied on the radial edges of the domain to force fluctuations to zero in these regions. These have a width of 10 cells (in the radial direction), both at the core boundary and the outer boundary. A particle source $S_n$ that is uniform in time and in the diamagnetic direction with Gaussian profile in the radial direction is applied. This source has its maximum 10 cells from the core boundary and has a standard deviation of 8 cells. In order to exclude the fringe regions and the region where the particle source is active, only the data on the central part of the domain (radially) is retained. Apart from that, the settings $\nu = D_0 = \chi_0$, $\Lambda = 2.8388$ are used in all simulations. The next three section detail the other parameters used in the simulations.

Appendix A.1. iso data set: isothermal SOL simulations
The isothermal SOL simulations are all conducted with a time step of $\Delta t = 1$ gyro period on a grid of 256x256 cells. All diamagnetic data points are used in the analysis but only for data in the radial range $51 \leq r \leq 199$. The electron temperature $T_e0$ is set to unity in all simulations. The parameters which are varied in these simulations are shown in table A1. The first simulation is the default one, only parameters that differ from this default simulation are shown for other simulations. This first simulation is also the one of used in figures 1 and 2.

| Nr. | $g$  | $T_e0$ | $\sigma$ | $\nu$ | Nr. | $g$  | $T_e0$ | $\sigma$ | $\nu$ |
|-----|------|--------|----------|-------|-----|------|--------|----------|-------|
| 1   | 6e-4 | 1      | 1e-4     | 5e-3  | 15  | 3.5e-4| 1      | 1e-4     | 5e-3  |
| 2   | 4e-4 | 16     | 1e-4     | 5e-3  | 16  | 4e-4 | 1      | 1e-4     | 5e-3  |
| 3   | 4.5e-4| 17     | 1e-4     | 5e-3  | 17  | 4e-4 | 1      | 1e-4     | 5e-3  |
| 4   | 7.5e-4| 18     | 1e-4     | 5e-3  | 18  | 9e-4 | 1      | 1e-4     | 5e-3  |
| 5   | 9e-4 | 19     | 1e-4     | 5e-3  | 19  | 8e-4 | 1      | 1e-4     | 5e-3  |
| 6   | 5e-5 | 20     | 1e-4     | 5e-3  | 20  | 9e-5 | 1      | 1e-4     | 5e-3  |
| 7   | 8e-5 | 21     | 1e-4     | 5e-3  | 21  | 4.5e-4| 1      | 1e-4     | 5e-3  |
| 8   | 2e-4 | 22     | 1e-4     | 5e-3  | 22  | 2e-4 | 1      | 1e-4     | 5e-3  |
| 9   | 0.5  | 23     | 1e-4     | 5e-3  | 23  | 9e-4 | 1      | 1e-4     | 5e-3  |
| 10  | 2    | 24     | 1e-4     | 5e-3  | 24  | 3e-4 | 4      | 1e-4     | 5e-3  |
| 11  | 2e-3 | 25     | 1e-4     | 5e-3  | 25  | 7e-4 | 4      | 1e-4     | 5e-3  |
| 12  | 4e-3 | 26     | 1e-4     | 5e-3  | 26  | 7e-4 | 4      | 1e-4     | 5e-3  |
| 13  | 6e-3 | 27     | 1e-4     | 5e-3  | 27  | 4e-4 | 4      | 1e-4     | 5e-3  |
| 14  | 1e-2 |        |          |       | 28  | 4e-4 | 4      | 1e-4     | 5e-3  |
Appendix A.2. isoDW data set: isothermal SOL+core region simulations

In the isothermal simulations with SOL and outer core region, a time step of $\Delta t = 1$ gyro period is used again and the electron temperature $T_e0$ is again set to unity in all simulations. The parameters which are varied in these simulations are shown in table A2, as well as the radial range $r_{\text{start}} \leq r \leq r_{\text{end}}$ used in the analysis. $L_x$ and $L_y$ denote the radial and diamagnetic domain length and $x_{\text{sep}}$ the location of the separatrix. Data from simulation 11 is used in figures 1 and 2.

Table A2. TOKAM2D and post-processing parameters of the simulations of the isoDW data set.

| Nr. | $g$   | $T_i0$ | $\sigma$ | $\sigma_{DW}$ | $\nu$ | $L_x$ | $L_y$ | $x_{\text{sep}}$ | $r_{\text{start}}$ | $r_{\text{end}}$ |
|-----|-------|--------|----------|---------------|-------|-------|-------|-----------------|-----------------|----------------|
| 1   | 2e-3  | 1      | 1.25e-4  | 5.9e-4        | 2.7e-3| 170   | 200   | 61             | 21              | 149            |
| 2   | 2.3e-4| 2.5e-5 | 2.74e-4  | 5e-3          |       | 170   | 200   | 61             | 21              | 149            |
| 3   | 6e-4  | 1e-4   | 1e-4     | 5e-3          |       | 170   | 200   | 61             | 21              | 149            |
| 4   | 2.3e-4| 3.6e-5 | 2.5e-5   | 2e-2          |       | 170   | 200   | 61             | 21              | 149            |
| 5   | 6e-4  | 1e-5   | 1e-4     | 5e-3          |       | 170   | 200   | 61             | 21              | 149            |
| 6   | 2.3e-4| 2.5e-5 | 2.74e-4  | 5e-3          |       | 256   | 256   | 105            | 21              | 149            |
| 7   | 6e-4  | 1e-4   | 1e-4     | 5e-3          |       | 256   | 256   | 105            | 21              | 149            |
| 8   | 2.3e-4| 3.6e-5 | 2.5e-5   | 2e-2          |       | 256   | 256   | 105            | 21              | 149            |
| 9   | 6e-4  | 1e-5   | 1e-4     | 5e-3          |       | 256   | 256   | 105            | 21              | 149            |
| 10  | 2.3e-4| 2.5e-5 | 2.74e-4  | 5e-3          |       | 384   | 256   | 129            | 21              | 339            |
| 11  | 6e-4  | 1e-4   | 1e-4     | 5e-3          |       | 384   | 256   | 129            | 21              | 339            |
| 12  | 2.3e-4| 3.6e-5 | 2.5e-5   | 2e-2          |       | 384   | 256   | 129            | 21              | 339            |
| 13  | 6e-4  | 1e-5   | 1e-4     | 5e-3          |       | 384   | 256   | 129            | 21              | 339            |

Appendix A.3. ani data set: anisothermal SOL simulations

The anisothermal SOL simulations are all conducted on a grid of 256x256 cells with a time step of $\Delta t = 0.1$ gyro period. All diamagnetic data points are used in the analysis but only for data in the radial range $51 \leq r \leq 199$. Electron and ion thermal energy sources $S_{pe}$ and $S_{pi}$ are applied that have the same shape as the particle source, they are determined as $S_p = 2/3T_{inj} S_n$. All simulations have been run with sheath heat transmission coefficients which are constant in space and in time with values $\gamma_i = 2.5$ and $\gamma_e = 4$. The parameters which are varied in these simulations are shown in table A3. The first simulation is the default one, only parameters that differ from this default simulation are shown for other simulations. This first simulation is also the one of used in figures 1 and 2.
Table A3. TOKAM2D parameters of the simulations of the $ani$ data set.

| Nr. | $g$   | $T_{inj,i}$ | $T_{inj,e}$ | $\sigma$ | $\nu$ | Nr. | $g$   | $T_{inj,i}$ | $T_{inj,e}$ | $\sigma$ | $\nu$ |
|-----|-------|-------------|-------------|----------|-------|-----|-------|-------------|-------------|----------|-------|
| 1   | 6e-4  | 3           | 3           | 1e-4     | 5e-3  | 13  | 8     | 2           | 2           | 6        | 6     |
| 2   | 2e-5  | 14          | 2           | 8        |       | 2   | 8     |             |             | 6        |       |
| 3   | 5e-5  | 15          | 2           | 2        |       | 16  | 13    | 6           | 6           | 2        |       |
| 4   | 8e-5  | 17          | 3           | 4        |       | 18  | 9     | 4           | 4           | 1.2e-3   |       |
| 5   | 1.5e-4| 19          | 1.2e-3      |          |       | 20  | 2.5e-3|             |             |          |       |
| 6   | 2e-4  | 21          |             | 1e-2     |       | 22  | 4e-4  |             |             |          |       |
| 10  | 5     | 23          | 8e-3        |          |       | 24  | 4     | 4           | 1.2e-4      |          |       |
| 11  | 6     | 2           | 4           | 4e-3     |       |     |       |             |             |          |       |
| 12  | 2     | 6           |             |          |       |     |       |             |             |          |       |
References

[1] Aho-Mantila L, Wischmeier M, Müller H, Potzel S, Coster D, Bonnin X, Conway G and the ASDEX Upgrade Team 2012 Nucl. Fusion 52 103006

[2] Reimold F, Wischmeier M, Bernert M, Potzel S, Coster D, Bonnin X, Reiter D, Meisl G, Kallenbach A, Aho-Mantila L, Stroth U and the ASDEX Upgrade Team 2015 J. Nucl. Mater. 463 128–134

[3] Dekeyser W, Bonnin X, Lisgo S W, PITTS R, Brunner D, LaBombard B and Terry J 2016 Plasma and Fusion Res. 11 140303

[4] Pope S B 2015 Turbulent Flows (Cambridge University Press)

[5] Bufferand H, Ciraolo G, Ph Ghendrih, Marandet Y, Bucalossi J, Colin C, Fedorczak N, Galassi D, Gunn J, Leybros R, Serre E and Tamin P 2016 Contrib. Plasma Phys. 56 555 – 562

[6] Coosemans R, Dekeyser W and Baelmans M 2021 Phys. Plasmas 28 012302

[7] Coosemans R, Dekeyser W and Baelmans M 2020 Contrib. Plasma Phys. 60 e201900156

[8] Sarazin Y and Ph Ghendrih 1998 Phys. Plasmas 5 4214–4228

[9] Marandet Y, Nace N, Valentinuzzi M, Tamin P, Bufferand H, Ciraolo G, Gensio P and Mellet N 2016 Plasma Phys. Contr. Fusion 58 114001

[10] Nace N 2018 Dynamics of driven and spontaneous transport barriers in the edge plasma of tokamaks Ph.D. thesis Aix-Marseille University available at http://www.theses.fr/2018AIXM0101

[11] De Wolf R 2020 Bayesian approach to parameter estimation and model validation for nuclear fusion reactor edge turbulence modelling Master’s thesis KU Leuven available through Lirias: 0706107_55418536

[12] De Wolf R, Coosemans R, Dekeyser W and Baelmans M Submitted to Nucl. Fusion

[13] Biglari H, Diamond P and Terry P 1990 Phys. Fluids B: Plasma Physics 2 1–4

[14] Hahm T and Burrell K 1995 Phys. Plasmas 2 1648

[15] Diamond P, Itoh S I, Itoh K and Hahm T 2005 Plasma Phys. Contr. Fusion 47 R35

[16] Manz P, Xu M, Fedorczak N, Thakur S and Tynan G 2012 Phys. Plasmas 19 012309

[17] Held M, Wiesenberger M, Kube R and Kendl A 2018 Nucl. Fusion 58 104001

[18] Nielsen A, Rasmussen J J, Olsen J, Leiten M, Hansen S, Yan N, Tophøj L and Wan B N 2017 Plasma Phys. Contr. Fusion 59 025012

[19] Baudoin C, Tamin P, Ciraolo G, Futtersack R, Gallo A, Ghendrih P, Marandet Y, Nace N and Norscini C 2016 Contrib. Plasma Phys. 56 563–568

[20] Baudoin C 2018 Numerical evaluations of mechanisms governing the heat transport in the edge plasma of tokamaks Ph.D. thesis Aix-Marseille University available at http://www.theses.fr/2018AIXM0053

[21] Canuto V M 1997 Astrophys. J. 482 827–851

[22] Scott B D 2003 Phys. Lett. A 320 53–62

[23] Scott B D 2005 New J. Phys. 7 92

[24] Trotta R 2008 Contemp. Phys. 49 71–104

[25] Knuth K, Habeck M, Malakar N, Mubeen A and Placek B 2015 Digit. Signal Process. 47 50–67

[26] Papadimitriou D and Papadimitriou C 2015 Comput Fluids 120 82–97

[27] Haario H, Saksman E and Tamminen J 2001 Bernoulli 7 223–242

[28] Gelman A, Roberts G and Gilks W 1996 Bayesian Statistics 5 599–607

[29] Chib S and Jeliazkov I 2001 J. Am. Stat. Assoc. 96 270–281

[30] Roy C and Oberkampf W 2010 48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition 092407

[31] Dose V, Neuhauser J, Kurzan B, Marmann H, Salzmann H and Upgrade Team A 2001 Nucl. Fusion 41 1671