13th International Symposium on Turbulence and Shear Flow Phenomena (TSFP13)
Montreal, Canada, June 25–28, 2024

SUBOPTIMAL TRANSIENT GROWTH: STATISTICAL AND DATA-DRIVEN VIEWPOINTS

Peter Frame
Department of Mechanical Engineering
University of Michigan
1350 Hayward Street, Ann Arbor, MI
pframe@umich.edu

Zhicheng Kai
Department of Mechanical Engineering
University of Michigan
1350 Hayward Street, Ann Arbor, MI
kaizz@umich.edu

Aaron Towne
Department of Mechanical Engineering
University of Michigan
1350 Hayward Street, Ann Arbor, MI
towne@umich.edu

ABSTRACT

We present two additions to the standard transient growth theory. The first is a statistical framework for transient growth that relates the statistics of the energy of evolving disturbances to the statistics of the initial disturbances. We use this framework to make a number of observations for Poiseuille flow, including that the mean growth is likely to be much smaller than the optimal growth and that the mean growth grows linearly with Reynolds number whereas the optimal growth grows quadratically. The second is a method for estimating the optimal growth from flow data, rather than the linearized Navier-Stokes operator.

Introduction

The standard approach for theoretical analysis of the stability of a steady-state solution of the Navier-Stokes equations is to linearize the equations around the solution and calculate the eigenvalues of the resulting linearized Navier-Stokes (LNS) operator. The steady-state solution is determined to be stable via this analysis if all the eigenvalues have a negative real part and unstable otherwise. This approach, usually referred to as modal stability theory, has many successes; predicting the critical Rayleigh number in Rayleigh-Bénard convection and the dominant wavenumber in the Kelvin-Helmholtz instability are two examples. The modal approach fails to predict, however, the instabilities that are observed in many shear flows. One dramatic example is that the eigenvalues around the parabolic profile in pipe flow are always negative regardless of the Reynolds number whereas pipe flow is observed to transition, usually in the range of $Re = 2000$. More broadly, the modal approach fails to predict the turbulence observed in Couette flow, plane Poiseuille flow, and boundary layer flows at certain conditions.

One mechanism leading to the turbulence observed in these shear flows is the potential for transient growth in such flows (Schmid & Henningson, 2001). Despite the fact that any disturbance in a (unforced) linear system with negative eigenvalues will decay eventually as $t \to \infty$, this decay may not be monotonic so long as the associated eigenfunctions are not orthogonal. Indeed, in shear flows, the eigenfunctions of the LNS operator are highly non-orthogonal, leaving a potential for large temporary growth of solutions to the linearized equations. Though the linear approximation is valid for the small initial disturbances, once they undergo large-scale growth, nonlinear effects may take over, leading to transition.

The importance of transient growth in a given flow is usually quantified by calculating the optimal growth, i.e., the greatest ratio of initial and final energy among all possible initial disturbances. This quantity, which we refer to as G^{opt}, can be calculated from the LNS operator $\mathbf{A}$ as the square of the largest singular value of the matrix exponential,

$$G^{opt}(t) = \sigma_1^2(\mathbf{L}e^{\mathbf{A}t}\mathbf{L}^{-1}) \tag{1}$$

where $\mathbf{L}$ is the Cholesky decomposition of the weight matrix used to quantify disturbance energy, $\mathbf{W} = \mathbf{L}^*\mathbf{L}$.

In this work, we describe two approaches beyond the standard transient growth analysis described above. The first is a statistical perspective (Frame & Towne, 2024) on growth motivated by the (rather obvious) insight that real disturbances to the flow will not exactly coincide with the optimal one. The question becomes whether, or under what conditions, near-optimal growth can be expected. We approach this question by deriving an equation for the mean energy of the disturbances at time t as a ratio of the mean initial energy of the disturbances, which we refer to henceforth as G^{mean}. This quantity depends on the statistics of the incoming disturbances – if these are likely to resemble the optimal disturbances, then the mean growth will be near the optimal growth, for example. By conducting numerical experiments with these initial statistics in plane Poiseuille flow, we make three key observations. First, the length scales in the initial disturbances have a great impact on the mean energy amplification, with shorter correlation lengths in the initial disturbances leading to substantially less growth. Second, while G^{opt} is known to scale quadratically with the Reynolds number, we find that G^{mean} scales only linearly in realistic cases. Finally, we find that the probability density function (PDF) of the growth is nearly a decaying exponential function, indicat-

ing that in cases where the $G^{mean}/G^{opt} \ll 1$, it is extremely unlikely for disturbances to undergo near-optimal growth.

Secondly, we present an approach to estimate G^{opt} from data. In the literature, G^{opt} is calculated exclusively from matrix operations on the LNS operator, and a data-driven approach is useful for a number of reasons. First, it allows estimation of G^{mean} in cases where the LNS operator is not available, e.g., in experiments or, in many cases, in simulation data. Second, in all but a few idealized cases, the LNS operator is too large to perform the matrix operations necessary to obtain G^{opt} analytically. A data-driven method lends itself to these cases as well because the LNS operator can be used to generate data from which G^{opt} can be estimated. We demonstrate the method by estimating the optimal spatial transient growth from data for a transitional boundary layer from the Johns Hopkins Turbulence Database (Li *et al.*, 2008), finding $G^{opt}(x)$ as well as the optimal output modes. Finally, we compare the G^{mean} and G^{opt} for this data, the former of which is easy to compute, finding that G^{opt}/G^{mean} with the values we obtain is more than 10^2 for this case.

Statistical framework

We first derive a formula for the mean energy amplification, finding that it depends on the correlation matrix of the initial disturbances. We then model this correlation matrix for Poiseuille flow and record a number of observations.

Mean energy amplification

A natural starting point for the statistical view is to derive the mean energy of disturbances at time t. This energy may be written as the inner product of the state at time t with itself $e(\boldsymbol{q}_t) = \boldsymbol{q}_t^* \mathbf{W} \boldsymbol{q}_t$, where $\mathbf{W}$ is a weight matrix. This energy may also be written as

$$e(\boldsymbol{q}_t) = \mathrm{tr}(\mathbf{L}\boldsymbol{q}_t\boldsymbol{q}_t^*\mathbf{L}^*) \tag{2}$$

where $\mathbf{L}^*\mathbf{L} = \mathbf{W}$ and where $\mathrm{tr}(\cdot)$ takes the trace of the argument. The evolved state is related to the initial one by an evolution operator $\boldsymbol{q}_t = \mathbf{M}_t\boldsymbol{q}_0$, so the energy at time t may be written in terms of the initial state as

$$e(\boldsymbol{q}_t) = \mathrm{tr}(\mathbf{L}\mathbf{M}_t\boldsymbol{q}_0\boldsymbol{q}_0^*\mathbf{M}_t^*\mathbf{L}^*) \tag{3}$$

Taking an expected value, we have

$$\mathbb{E}[e(\boldsymbol{q}_t)] = \mathrm{tr}(\mathbf{L}\mathbf{M}_t\mathbf{C}_{00}\mathbf{M}_t^*\mathbf{L}^*) \tag{4}$$

where $\mathbf{C}_{00} = \mathbb{E}[\boldsymbol{q}_0\boldsymbol{q}_0^*]$ is the correlation matrix of the initial disturbances. Finally, to quantify the growth (rather than that energy) we divide by the expected initial energy

$$G^{mean} = \frac{\mathrm{tr}(\mathbf{L}\mathbf{M}_t\mathbf{C}_{00}\mathbf{M}_t^*\mathbf{L}^*)}{\mathrm{tr}(\mathbf{L}\mathbf{C}_{00}\mathbf{L}^*)} \tag{5}$$

The quantity G^{mean} is the ratio of the expected energy at time t to the expected initial energy, i.e., it is the amount by which the mean energy is amplified over the time interval. More generally, the correlation at time t is related to the initial correlation as

$$\mathbf{C}_{tt} = \mathbf{M}_t\mathbf{C}_{00}\mathbf{M}_t^* \tag{6}$$

Perhaps unsurprisingly, we see that the second-order statistics, i.e., the mean and correlation, of the disturbances at time t depend on the second-order statistics of the initial disturbances. The initial correlation must, therefore, be modeled to get a numerical value for G^{mean}.

Application to plane Poiseuille flow

We apply the framework described above to plane Poiseuille flow. The numerics are handled with the code provided in Schmid & Henningson (2001), which employs a Chebyshev discretization of the linearized Navier-Stokes equations for channel flow (Herbert, 1977; Reddy & Henningson, 1993) and uses the wall-normal velocity v and vorticity η as coordinates. A number of observations are apparent with the statistical lens.

Most transient growth analyses consider disturbances at particular streamwise and spanwise wavenumbers, α and β, respectively, as opposed to disturbances that have more general streamwise and spanwise dependence. We take this as our starting point as well, but show later that this choice alone leads to substantial exaggeration of the mean growth. We take the initial correlation $\mathbf{C}_{00}$ to contain energy only in v (which promotes growth), and to be Gaussian-correlated in the wall-normal direction, i.e.,

$$\mathbf{C}_{00}^{\eta\eta}(y_1, y_2, \alpha, \beta) = \exp[-\left(\frac{y_1 - y_2}{\lambda}\right)^2] \tag{7}$$

Here, λ is the correlation length in the wall-normal direction at $Re = 1000$. The resulting $G^{mean}(t)$ optimized over time, which we refer to as G^{mean}_{max}, is shown in Figure 1. At all of the

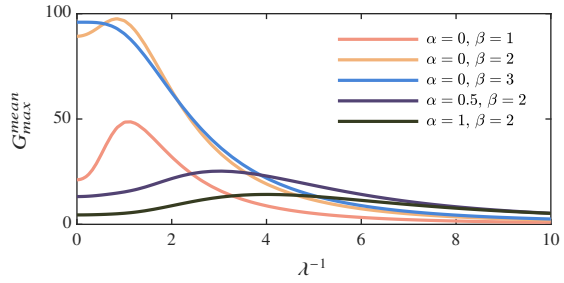

Figure 1: G^{mean}_{max} as a function of the inverse of the correlation length λ^{-1} in Poiseuille flow. The expected growth decreases as the correlation length is shortened.

wavenumber pairs shown, short structures in the wall-normal direction (large λ^{-1}) grow by less, and there is a non-zero peak. The optimal growth over all wavenumbers for this flow at $Re = 1000$ is roughly 200 (Schmid & Henningson, 2001), thus G^{mean}_{max} for $(\alpha, \beta) = (0, 2)$ is comparable in scale to the optimal growth when the wall-normal correlation length is long.

However, just as the y-dependence of the disturbances to the flow will not happen to coincide with that of the optimal one, the x and z dependence of real disturbances will not be optimal either. That is, disturbances will not be perfectly streamwise-

and spanwise-periodic with the optimal wavenumbers. The impact of this additional nonoptimality can be assessed by considering three-dimensional initial correlation matrices, and we find this, again, leads to substantially less growth.

We take the correlation to be

$$\mathbf{C}_{00}^{\eta\eta}(\boldsymbol{x}_1,\boldsymbol{x_2}) = \exp\left[-\left(\frac{|\boldsymbol{x}_2-\boldsymbol{x}_1|}{\lambda}\right)^2\right] \tag{8}$$

where $\boldsymbol{x}_1 = [x_1, y_1, z_1]^T$, and where $|\cdot|$ returns the Euclidean length of the argument. In other words, (8) is an isotropic Gaussian with correlation length λ with all its energy in the vorticity. The growth G_{max}^{mean} that results is shown as a function of Reynolds number in Figure 2 for various choices of λ, along with G_{max}^{opt}. Two aspects are notable. First, at $Re = 1000$, $G_{max}^{mean} = 4.5$ for $\lambda = 1$ (the maximum over all λ occurs for $\lambda = 0.83$, and gives a similar G_{max}^{mean}). Thus, G_{max}^{mean} is only 2.5% of G_{max}^{opt} at the same Reynolds number. Second, while G_{max}^{opt} scales quadratically with Re, G_{max}^{mean} scales only linearly. This means that the amount by which G^{opt} overstates the growth of realistic disturbances increases with Reynolds number.

Does the fact that the mean energy of disturbances is far smaller than the maximum preclude the possibility of occasional large-growth events? In other words, what is the probability of achieving large multiples of the mean? To answer questions such as these, the probability density function (PDF) of the disturbances at time t must be obtained, and to do this, the PDF of the initial disturbances must be modeled. In Figure 3, we show the results of two different initial disturbance PDFs, each with the same correlation. The first is a multivariate Gaussian (MVG) distribution, and the second is the uniform distribution on the N-sphere transformed linearly so that it has the desired correlation (see Frame & Towne (2024) for details). That both distributions have the same correlation implies that the mean energies are the same.

First, the two empirical energy PDFs are quite similar despite being generated by different initial disturbance PDFs. This is an indication that, although the distribution of energy does formally depend on the details of the distribution of the incoming disturbances, practically this dependence is quite weak. This allows for an approximation of the distribution using only information about the incoming correlation $\mathbf{C}_{00}$. The approximation shown is obtained by analytically calculating the fourth moment of the energy PDF generated by the MVG distribution and finding the exponential distribution that shares this fourth moment.

Data-driven method for calculating optimal transient growth

Despite the fact that transient growth is an important mechanism for transition in many shear flows, its precise influence is not easily assessed in many cases. Calculating $G^{opt}(t)$ involves computing $\exp[\mathbf{A}t]$ with $\mathbf{A} \in \mathbb{C}^{N_x \times N_x}$. Perhaps even more prohibitive for many flows of interest is that obtaining $\mathbf{A}$ itself requires linearizing the governing equations around the base flow; even with access to the governing equations, this linearization can be remarkably cumbersome. Thus, it is useful to have a non-intrusive means of calculating G^{opt}, i.e., a method for taking simulation or experiment data and estimating G^{opt}.

Formulation

We assume access to a set of initial disturbances $\{\boldsymbol{q}_0^1, \boldsymbol{q}_0^2, \ldots, \boldsymbol{q}_0^m\}$ and to the same disturbances some time t later, $\{\boldsymbol{q}_t^1, \boldsymbol{q}_t^2, \ldots, \boldsymbol{q}_t^m\}$. These disturbances are related by

$$\boldsymbol{q}_t^i = \mathbf{M}_t \boldsymbol{q}_0^i \tag{9}$$

for some unknown operator $\mathbf{M}_t$, which is assumed to be linear. The growth undergone by the i-th disturbance is $\|\boldsymbol{q}_t^i\|^2/\|\boldsymbol{q}_0^i\|^2$, but this is unlikely to be the maximum growth possible in the system. With the assumption that the evolution operator is linear, we can determine the evolution of any vector in the span of the initial disturbances as

$$\mathbf{M}_t\left(\sum_{j=1}^{m} \boldsymbol{q}_0^j \psi_j\right) = \sum_{j=1}^{m} \boldsymbol{q}_t^j \psi_j \tag{10}$$

where the ψ_j are scalar coefficients. The optimal growth is approximated by finding the linear combination of the initial and evolved states that corresponds to the largest growth. This may be written formally as the following optimization problem for the coefficients $\boldsymbol{\psi}$,

$$G_{dd}^{opt} = \max_{\boldsymbol{\psi}} \frac{\|\mathbf{Q}_t\boldsymbol{\psi}\|^2}{\|\mathbf{Q}_0\boldsymbol{\psi}\|^2} \tag{11}$$

where the data matrices are $\mathbf{Q}_0 = [\boldsymbol{q}_0^1, \boldsymbol{q}_0^2, \ldots, \boldsymbol{q}_0^m]$, and $\mathbf{Q}_t = [\boldsymbol{q}_t^1, \boldsymbol{q}_t^2, \ldots, \boldsymbol{q}_t^m]$, and G_{dd}^{opt} is the data-driven approximation of the optimal transient growth. As before, the weight matrix $\mathbf{W}$ is used to define the inner product and thus the norm. The optimization above may be written with $\mathbf{L}$, the Cholesky decomposition of the weight matrix, as

$$G_{dd}^{opt} = \max_{\boldsymbol{\psi}} \frac{\boldsymbol{\psi}^*\mathbf{Q}_t^*\mathbf{L}^*\mathbf{L}\mathbf{Q}_t\boldsymbol{\psi}}{\boldsymbol{\psi}^*\mathbf{Q}_0^*\mathbf{L}^*\mathbf{L}\mathbf{Q}_0\boldsymbol{\psi}} \tag{12}$$

This may be written as a Rayleigh quotient by introducing the transformation $\boldsymbol{\nu} = \mathbf{B}\boldsymbol{\psi}$, with $\boldsymbol{B}$ defined by $\mathbf{B}^*\mathbf{B} = \mathbf{Q}_0^*\mathbf{L}^*\mathbf{L}\mathbf{Q}_0$. With this transformation, (12) becomes

$$G_{dd}^{opt} = \max_{\boldsymbol{\nu}} \frac{\boldsymbol{\nu}^*\mathbf{B}^{-1*}\mathbf{Q}_t^*\mathbf{L}^*\mathbf{L}\mathbf{Q}_t\mathbf{B}^{-1}\boldsymbol{\nu}}{\boldsymbol{\nu}^*\boldsymbol{\nu}} \tag{13}$$

and the Rayleigh quotient is optimized when $\boldsymbol{\nu} = \boldsymbol{u}_1(\mathbf{L}\mathbf{Q}_t\mathbf{B}^{-1})$, where $\boldsymbol{u}_1(\cdot)$ returns the first left singular vector of the argument. The optimizing coefficients are

$$\boldsymbol{\psi} = \boldsymbol{B}^{-1}\boldsymbol{u}_1(\mathbf{L}\mathbf{Q}_t\mathbf{B}^{-1}) \tag{14}$$

and the data-driven approximation of the optimal growth is

$$G_{dd}^{opt} = \sigma_1^2(\mathbf{L}\mathbf{Q}_t\mathbf{B}^{-1}) \tag{15}$$

where $\sigma_1(\cdot)$ returns the first singular value of the argument. We note that in practice, the method must be regularized,

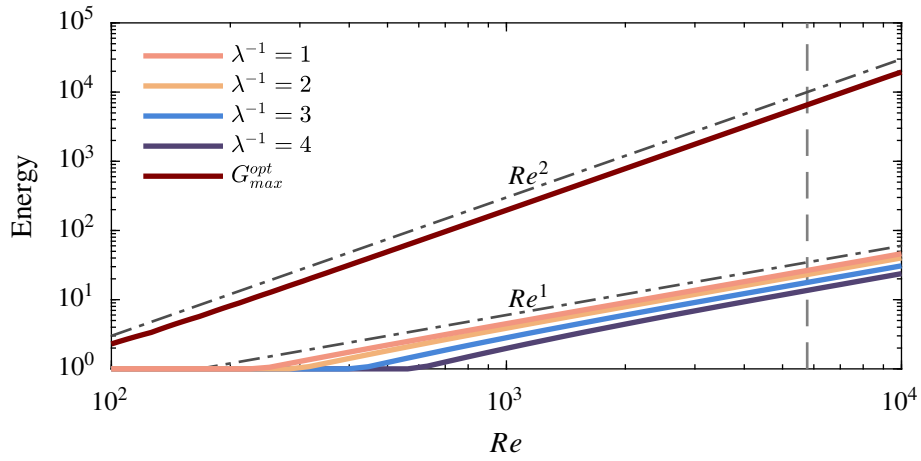

Figure 2: Linear scaling of G^{mean}_{max} is observed in contrast to the well-known quadratic scaling of G^{opt}_{max}. Here the initial correlations are isotropic and the correlation between two points is a Gaussian function of their three-dimensional separation with correlation length λ. This mixes wavenumbers which leads to far less growth.

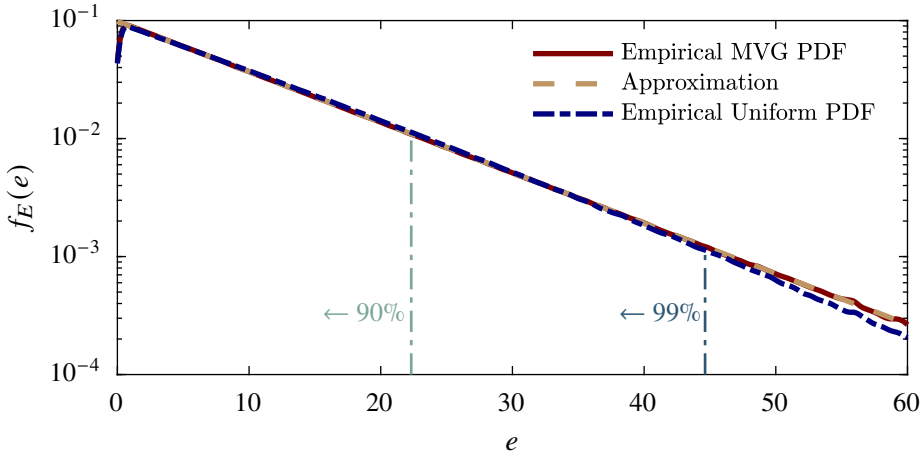

Figure 3: The empirically calculated energy PDFs for the multivariate Gaussian initial disturbance PDF and transformed uniform PDF. Both result in very similar PDFs for the energy, which are nearly exponential. The approximation is calculated without the Monte-Carlo and is accurate for both.

which is done by modifying the denominator in (12) to be $\boldsymbol{\psi}^*(\mathbf{Q}_0^*\mathbf{L}^*\mathbf{L}\mathbf{Q}_0 + m\gamma\mathbf{I})\boldsymbol{\psi}$, where γ is a regularization parameter.

Connection to dynamic mode decomposition

The method may be equivalently formulated as approximating the linear operator $\mathbf{M}_t$ with a similar procedure to that used in dynamic mode decomposition (DMD) (Schmid, 2010), then taking the first (square) singular value of this approximation. To properly account for the weight, we work with the weighted states $\mathbf{L}\boldsymbol{q}$. Analogous to DMD, we look for the matrix that most closely (in the least squares sense) maps the initial (weighted) states to the evolved ones,

$$\mathbf{L}\mathbf{Q}_t \approx \mathbf{M}_{t,dd}\mathbf{L}\mathbf{Q}_0 \tag{16}$$

This operator is obtained by taking the pseudoinverse

$$\mathbf{M}_{t,dd} = \mathbf{L}\mathbf{Q}_t(\mathbf{L}\mathbf{Q}_0)^+ = \mathbf{L}\mathbf{Q}_t\mathbf{B}^{-1}\mathbf{B}^{-1*}\mathbf{Q}_0^*\mathbf{L}^* \tag{17}$$

The first square singular value of this matrix gives the data-driven approximation of G^{opt},

$$G^{opt}_{dd} = \sigma_1^2(\mathbf{L}\mathbf{Q}_t\mathbf{B}^{-1}\mathbf{B}^{-1*}\mathbf{Q}_0^*\mathbf{L}^*) = \sigma_1^2(\mathbf{L}\mathbf{Q}_t\mathbf{B}^{-1}) \tag{18}$$

The second equality shows the DMD formulation is equivalent to the maximization formulation. It holds because 1) the singular values of any matrix are equivalent to the eigenvalues of the matrix multiplied by its conjugate transpose and 2) $\mathbf{B}^{-1*}\mathbf{Q}_0^*\mathbf{L}^*$ multiplied by its conjugate transpose is the identity.

Recently, a similar objective of estimating transient growth from data was pursued by Dotto *et al.* (2022). Our approach differs in one important way from this work: Specifically, Dotto *et al.* (2022) found the initial disturbance with the largest growth rate at $t = 0$, then evolved this disturbance forward in time, computing its growth along the way. However, this is not the optimal growth G^{opt} – the optimal growth is not the evolution curve of any particular disturbance. Rather, it is the envelope of all initial disturbance curves, and the method we have described above reflects this fact.

Application to transitional boundary layer

We apply the data-driven method described above to transitional boundary layer data from the Johns Hopkins Turbulence Data Base (Li *et al.*, 2008; Zaki, 2013). We show, by applying the data-driven algorithm, that while very large growth is possible, it is not realized and G^{mean} is a rather modest 1.4 when integrated over β and ω. The transitional boundary layer lends itself to a spatial, rather than temporal, transient growth analysis. We briefly introduce this concept next, then describe the boundary layer data and results of application for the data-driven method to it.

Spatial transient growth So far, we have discussed transient growth within a temporal framework, i.e., the flow is disturbed at a particular time $t = 0$, and the disturbance evolves through time. Transient growth may also be studied in a spatial framework wherein disturbances occur at $x = 0$, then evolve downstream without further forcing to the equations. Whereas in the temporal context, the governing equations were transformed in x and z corresponding to α and β, now the equations are transformed in t and z. Thus, we have a temporal frequency ω in addition to the spanwise wavenumber β. The governing equations are now taken to be (Herbert, 1997; Towne & Colonius, 2015; Zhu & Towne, 2023)

$$\frac{\partial}{\partial x}\boldsymbol{q}(x,\beta,\omega) = \mathbf{A}(x,\beta,\omega)\boldsymbol{q}(x,\beta,\omega) \tag{19}$$

In a parallel flow, $\mathbf{A}$ is independent of x, but this dependence must be retained in nonparallel flows. The operator to be approximated $\mathbf{M}_x(\beta,\omega)$ maps $\boldsymbol{q}(x=0,\beta,\omega)$ to $\boldsymbol{q}(x,\beta,\omega)$.

Data description and computational considerations The data we use are from a DNS of a transitional boundary layer over a flat plate with a semicircular leading edge and plate thickness L_p (Zaki, 2013). The domain to which we have access begins a small streamwise distance from the beginning of the boundary layer and extends well into the turbulent region in x and the free stream in y. It is periodic in the spanwise direction. The data consist of 4701 snapshots of dimension $n_x = 3320$, $n_y = 224$, $n_z = 2048$. We downsample by a factor of 2 in all spatial directions and limit the calculations to the non-turbulent part of the domain because the method, which is built on the assumption of linear evolution, breaks down once nonlinearity becomes important to the dynamics. The inner product, defined by $\mathbf{W}$, is based on the kinetic energy of disturbances.

To build the data matrices needed for the algorithm, many realizations of the flow at the x, β, and ω values of interest are needed. To generate multiple realizations at the same temporal frequency ω, the data, a single long time series, must be partitioned into m (possibly overlapping) smaller time series (Welch, 1967). The temporal FFT of each block is then taken, and the result at each frequency is stored as one realization in the corresponding data matrix (see Towne *et al.* (2018) for a schematic).

Optimal growth and output mode The results of the data-driven method at $\omega = 0$ are shown for three values of $\beta\delta_{x_0}$ in Figures 4 and 5, where δ_{x_0} is the displacement thickness of a Blasius profile at the initial x-location. The optimal growth, shown in Figure 4, is quite large for all three $\beta\delta_{x_0}$, and as is expected, structures at nonzero β are capable of more growth than those at $\beta = 0$. The $\beta\delta_{x_0} = 0.263$ curve reaches a local maximum at $Re_x = 5.5\times10^4$, though all three curves attain higher values outside of the domain plotted. We believe this is due to nonlinearities becoming prominent downstream of the streamwise domain shown. The results for the growth are somewhat preliminary – they are sensitive to the choice of the regularization parameter γ, and further investigation is required to remove this sensitivity.

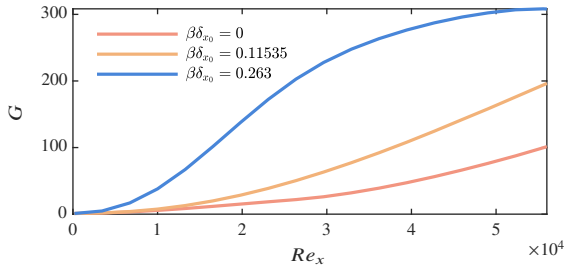

Figure 4: G^{opt} obtained from the data-driven method for three wavenumbers at $\omega = 0$. There is a potential for large-scale growth, and the largest occurs for finite β.

In Figure 5, the streamwise velocity components of the output modes are shown for the three values of $\beta\delta_{x_0}$ at $Re_x = 5.8\times10^4$. Here, y is scaled by the local displacement thickness at the output location. The structures are very similar for the two non-zero values of β and match the shape of those in Andersson *et al.* (1999) and Luchini (2000). For the $\beta = 0$ case, the mode resembles a Tollmien-Schlichting wave (Schmid & Henningson, 2001), which is likely the fastest growing mode at $\beta = 0$. Also, we note that these structures are insensitive to γ, unlike the G^{opt} results.

Comparison of G^{mean} and G^{opt} for the boundary layer The rather large growth reported above can be compared to G^{mean} for the boundary layer. It is relatively straightforward to calculate G^{mean} given data — G^{mean} is simply the ratio of the average energies at two streamwise locations in the flow. The average can be computed as a function of β, or can be averaged over β. Specifically G^{mean} as a function of $\beta\delta_{x_0}$ is

$$G^{mean}(x,\beta) = \frac{\mathbb{E}[\|\boldsymbol{q}(x,\beta)\|^2]}{\mathbb{E}[\|\boldsymbol{q}(0,\beta)\|^2]} \tag{20}$$

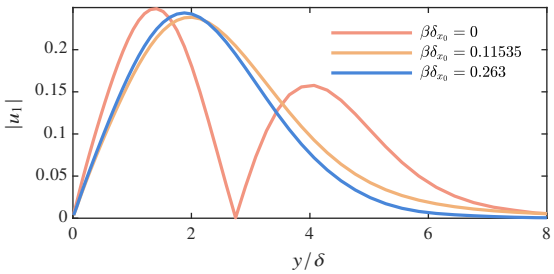

Figure 5: The output modes u_1 at $Re_x = 5.8\times10^4$, which is near the peak x-location for $\beta\delta_{x_0} = .263$

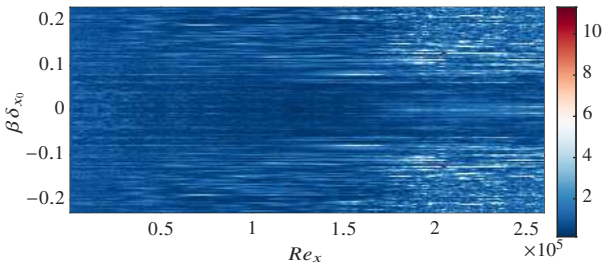

Figure 6: Mean energy amplification as a function of $\beta\delta_{x_0}$ and x. While very large growth is possible (see Figure 4), little growth is realized on average at any wavenumbers.

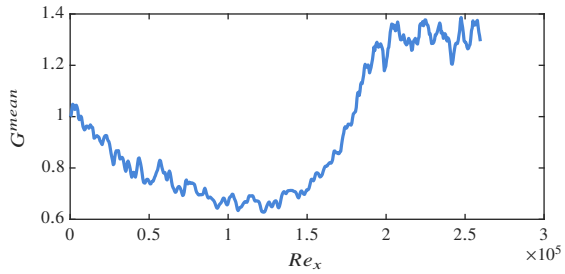

Figure 7: Mean energy amplification as a function of x. There is little growth, on average, despite the large potential for it.

and as a function of just x is

$$G^{mean}(x) = \frac{\mathbb{E}[\int_{-\infty}^{\infty} \|\boldsymbol{q}(x,\beta)\|^2 \, d\beta]}{\mathbb{E}[\int_{-\infty}^{\infty} \|\boldsymbol{q}(0,\beta)\|^2 \, d\beta]} \tag{21}$$

$G^{mean}(x,\beta\delta_{x_0})$ is shown in Figure 6. The largest values shown are for x-locations where the nonlinearity is substantial, and the largest $G^{mean}(x,\beta\delta_{x_0})$ is orders of magnitude lower than G^{opt} (regardless of the regularization parameter). Figure 7 shows $G^{mean}(x)$, i.e., the turbulent kinetic energy (TKE) ratio of the flow at streamwise locations x and 0. The TKE drops initially before ascending just above unity, further underscoring the fact that the optimal growth can vastly overpredict the mean growth. An initial drop in G^{mean} was reported in Frame & Towne (2024) for all initial correlations used, but more work would be required to determine whether the initial drop here is due to the same mechanism.

Conclusions

We have introduced two augmentations to the standard transient growth theory. The first is a statistical lens through which to view transient growth that goes beyond the standard optimal-disturbance approach, and the second is a method for determining the level of optimal transient growth without direct access to the LNS operator using flow data. We demonstrated the statistical framework on Poiseuille flow and used it to show that 1) long structures in the wall-normal direction tend to grow by more, 2) the mean growth scales linearly with Reynolds number in contrast with the quadratic scaling of the optimal growth, and 3) that the PDF of the growth is nearly exponential, meaning that large-growth events (ones more than a small multiple of the mean) are quite rare. More broadly, we showed that, when various factors are taken into account, the level of transient growth is likely to be much smaller than the optimal value. We used the data-driven method to calculate the optimal growth and optimal output modes for a transitional boundary layer. The optimal growth structures were similar in shape to those found in Andersson *et al.* (1999) and Luchini (2000), and the optimal growth level was on the order of 10^2, though there is uncertainty in the growth result due to dependence on the regularization parameter. We also calculated the mean energy amplification for the same flow and found it to be order unity, supporting our view that the optimal growth is an insufficient metric for quantifying the importance of transient growth.

REFERENCES

Andersson, Paul, Berggren, Martin & Henningson, Dan S. 1999 Optimal disturbances and bypass transition in boundary layers. *Physics of Fluids* **11** (1), 134–150.

Dotto, A., Barsi, D., Lengani, D. & Simoni, D. 2022 A data-driven optimal disturbance procedure for free-stream turbulence induced transition. *Physics of Fluids* **34** (12).

Frame, Peter & Towne, Aaron 2024 Beyond optimal disturbances: a statistical framework for transient growth. *Journal of Fluid Mechanics* **983**.

Herbert, Thorwald 1977 Die neutrale fläche der ebenen Poiseuille-Strömung. habilitation, Uni Stuttgart.

Herbert, Thorwald 1997 Parabolized stability equations. *Annual Review of Fluid Mechanics* **29** (1), 245–283.

Li, Yi, Perlman, Eric, Wan, Minping, Yang, Yunke, Meneveau, Charles, Burns, Randal, Chen, Shiyi, Szalay, Alexander & Eyink, Gregory 2008 A public turbulence database cluster and applications to study lagrangian evolution of velocity increments in turbulence. *Journal of Turbulence* **9**, N31.

Luchini, PAOLO 2000 Reynolds-number-independent instability of the boundary layer over a flat surface: optimal perturbations. *Journal of Fluid Mechanics* **404**, 289–309.

Reddy, Satish C. & Henningson, Dan S. 1993 Energy growth in viscous channel flows. *J. Fluid Mech.* **252**, 209–238.

Schmid, Peter & Henningson, Dan 2001 *Stability and Transition in Shear Flows*. Springer.

Schmid, Peter J. 2010 Dynamic mode decomposition of numerical and experimental data. *Journal of Fluid Mechanics* **656**, 5–28.

Towne, Aaron & Colonius, Tim 2015 One-way spatial integration of hyperbolic equations. *Journal of Computational Physics* **300**, 844–861.

Towne, Aaron, Schmidt, Oliver T. & Colonius, Tim 2018 Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *Journal of Fluid Mechanics* **847**, 821–867.

Welch, P. 1967 The use of fast fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics* **15** (2), 70–73.

Zaki, Tamer A. 2013 From streaks to spots and on to turbulence: Exploring the dynamics of boundary layer transition. *Flow, Turbulence and Combustion* **91** (3), 451–473.

Zhu, Min & Towne, Aaron 2023 Recursive one-way navier-stokes equations with pse-like cost. *Journal of Computational Physics* **473**, 111744.