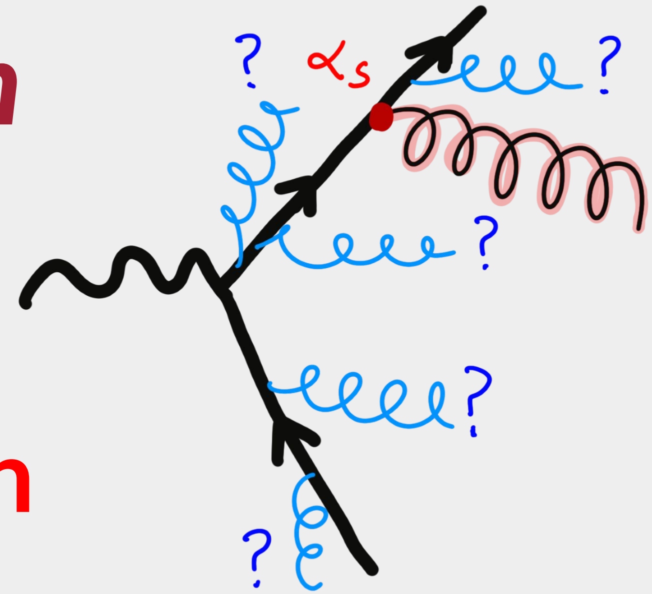
[RG, Mastandrea; 2508.XXXX] — Coming soon!

Title Change!

~~Renormalizing Flows:~~

Resummed Distribution Functions:

Unitary Completions of Fixed-Order Perturbation Theory

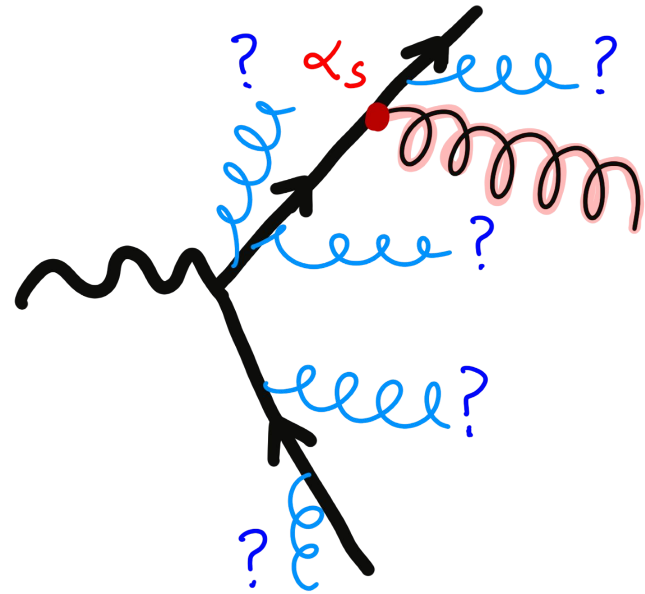


Rikab Gambhir

In Collaboration with: Radha Mastandrea

Basic Premise

FACT 1: Rikab is not very good at hardcore QCD calculations. Resummation is *hard*!

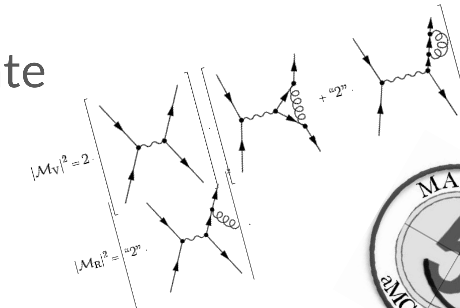


Basic Premise

FACT 1: Rikab is not very good at hardcore QCD calculations. Resummation is *hard*!

Large logs? IR Divergences? Resummation? SCET? Plus functions? Sudakov Shoulders? Multiple scales? Too much for me!
Can I just guess?

Suppose I can calculate the *fixed-order* distribution for observable(s) *x*:



$$\underbrace{p(x|\alpha) = p_0(x) + \alpha p_1(x|\alpha) + \alpha^2 p_2(x|\alpha) + ... + \alpha^M p_M(x|\alpha)}_{\text{Fixed-order calculation}} + \underbrace{\mathcal{O}\left(\alpha^{M+1}\right)}_{\textit{Unknown}}$$

FACT 2: He at least knows what a probability distribution *should* look like.

How far can we get with just *guessing* and enforcing consistency conditions?
What is the full distribution even *allowed* to be?

*By "smooth", we mean comprised of strict functions

Basic Premise – Continued

$$\underbrace{p(x|\alpha) = p_0(x) + \alpha p_1(x|\alpha) + \alpha^2 p_2(x|\alpha) + \ldots + \alpha^M p_M(x|\alpha)}_{\text{Known Fixed-order calculation}} + \underbrace{\mathcal{O}\left(\alpha^{M+1}\right)}_{\text{Unknown}}$$

Goal: Use consistency conditions, *and nothing else*, to say *something* about this.

Consistency Conditions

The full, all-orders distribution (nature) should be:

1. Positive
2. Finite e.g. no "large logs"
3. Normalized
4. Smooth*

A fixed-order calculation is typically *none* of these things. Let's enforce it!

Our Strategy:

Ansatz
The Resummed Distribution Function

$\xrightarrow{\text{Perturbative Info}}$

Result
All non-pathological distributions consistent with perturbative series.

1. Write down all *possible* consistent distributions, with free parameters

2. Fix some of the free parameters with perturbative info

3. Remaining free parameters parameterize all allowed higher order terms!

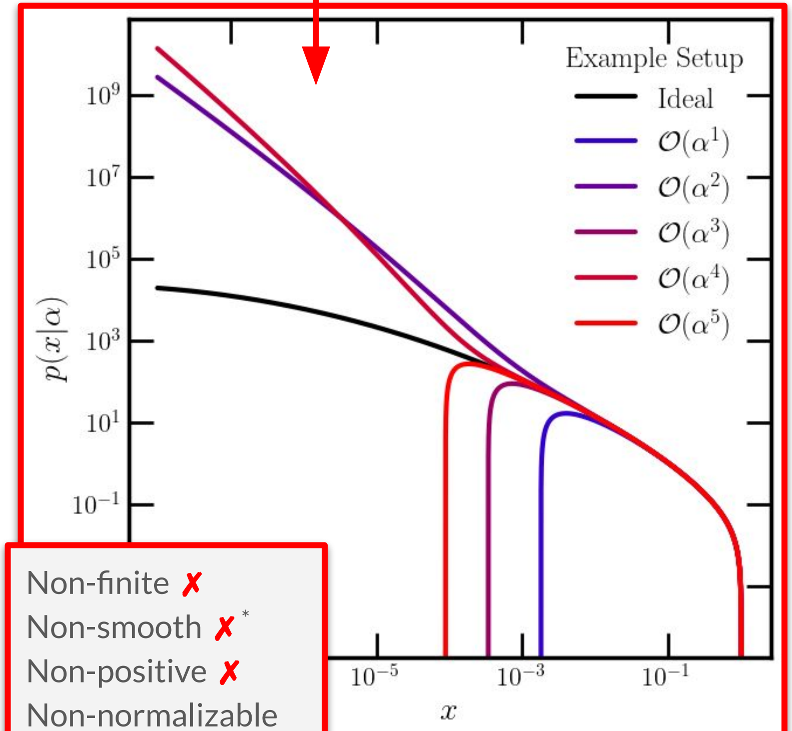
*(not here, but e.g. C param and similar kinematic bounds)

Basic Premise – Spelled Out

$$p(x|\alpha) = \underbrace{p_0(x) + \alpha p_1(x|\alpha) + \alpha^2 p_2(x|\alpha) + \ldots + \alpha^M p_M(x|\alpha)}_{\text{Known Fixed-order calculation}} + \underbrace{\mathcal{O}\left(\alpha^{M+1}\right)}_{\text{Unknown}}$$

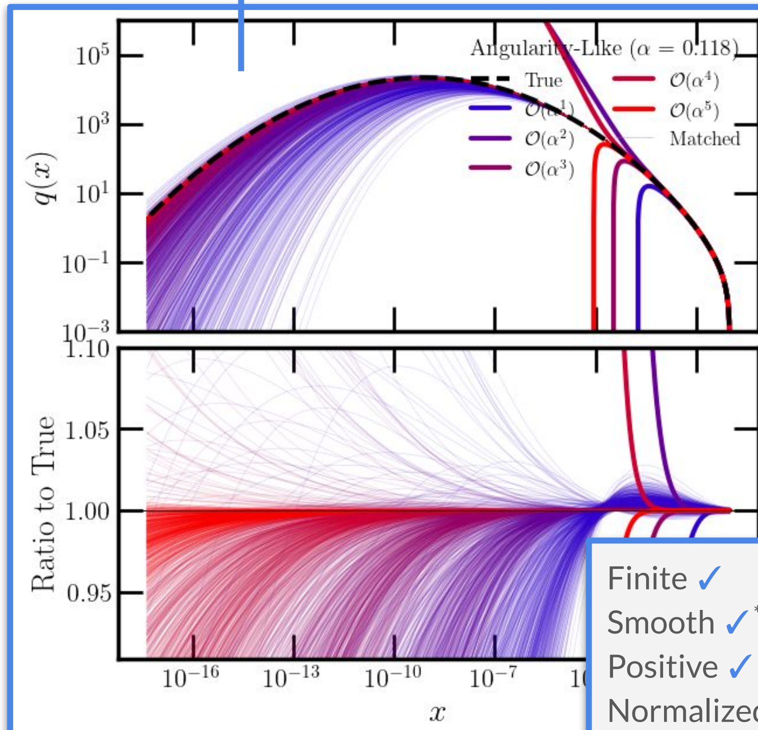
You give us this

We give you back this



Non-finite ✗
Non-smooth ✗*
Non-positive ✗
Non-normalizable ✗

Matching → To RDF



Finite ✓
Smooth ✓*
Positive ✓
Normalized ✓

Fixed-order distribution p in α for an observable x

All non-pathological all-orders distributions q that match p to the given order

The Resummed Distribution Function (RDF)

Given a set of random variables t_i valued over the positive reals conditioned on a perturbative parameter α, we define the **Resummed Distribution Function (RDF)** q

$$q(t_1, t_2, ..., t_k|\alpha) = f_1(t_1, \alpha) \times ... \times f_k(t_1, t_2, ..., t_k, \alpha)$$
$$\times \exp\left[-\int_0^{t_1} dt_1'\, f_1(t_1', \alpha)\right] \times ... \times \exp\left[-\int_0^{t_k} dt_k'\, f_k(t_1, t_2, ..., t_k', \alpha)\right]$$

where $\quad f_j(t_1, t_2, ..., t_j, \alpha) = \exp[-g_j(t_1, t_2, ..., t_j, \alpha)]$

This ansatz parametrizes *all* valid conditional PDFs that are everywhere finite, positive, normalized, and analytic in α.

The ansatz is specified by the functions g, which are only mildly-constrained functions and can be almost *anything* (neural networks, polynomials, …)

The Resummed Distribution Function [Details]

Given a set of random variables t_i valued ... d on a perturbative parameter α, we define the ... q

For an observable x from 0 to 1, we transform to t from 0 to ∞

$$t \equiv \log\left(\frac{1}{x}\right)$$

This is where logs come from!

$$\ldots_k(t_1, t_2, ..., t_k, \alpha)$$

$$\ldots(t_1', \alpha)\Big] \times ... \times \exp\left[-\int_0^{t_k} dt_k'\, f_k(t_1, t_2, ..., t_k', \alpha)\right]$$

$$\ldots t_2, ..., t_j, \alpha) = \exp[-g_j(t_1, t_2, ..., t_j, \alpha)]$$

... ditional PDFs th ...

Assumptions

We assume the expansion $p_m(x)$ is the fixed-order expansion of *some* all-orders distribution – *not always true!*

We can only capture perturbative information. No non-perturbative info!

These conditions can be violated in QCD (asymptotic series, etc). Be careful!

The only requirements on g are that it does not diverge to positive infinity, and that it is analytic in α up to single logs

The ansatz is specified by the functions g, which are only mildly-constrained functions and can be almost *anything* (neural networks, polynomials, ...)

We choose polynomials in t and α since this is related to logarithmic resummation

$$g(t, \alpha) = \sum_{m=0,n=1}^{M,N} g_{mn}\alpha^m t^n \Theta(t - \theta_{mn}) + \log(g^*(t, \alpha))$$

[Aside] Connection to Logarithmic Resummation

The RDF is *an* all-orders resummation that matches the F.O. calculation. It is not *the* logarithmic resummation — but they relate!

$$P(x|\alpha) = (1 + \sum_{m=1}^{\infty} C_m \alpha^m) \times \exp\left[-\sum_{m=1}^{\infty}\sum_{n=1}^{m+1} F_{mn}\alpha^m \log^n(1/x)\right] + \cancel{R(x)}$$

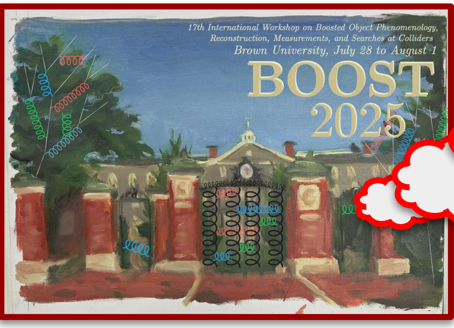
Non-Singular

Leading Log: All terms $F_{n,n+1}$ included
Next-to-Leading Log: All terms $F_{n,n}$ included

Maps directly onto the RDF *only if* (a) *f* is polynomial and *(b)* all relevant logs are included in the F.O. calculation to match (including running coupling, NGLs, etc)

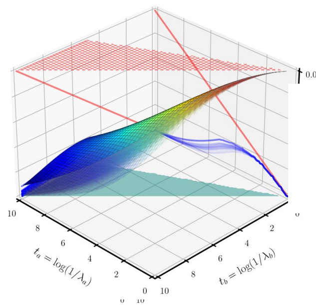
$$Q(x|\alpha) = \exp[-F(\log(1/x), \alpha)], \quad \text{where } F(t,\alpha) = \int_0^t dt' f(t,\alpha) \qquad f(t,\alpha) = \exp[-g(t,\alpha)]$$

Higher order unconstrained *g* = "guesses" of the $N^{n>N}LL$ logarithms, can set to zero



Why ever do this?

1. If you don't have access to the logarithmic resummation, build *informed guesses*
   a. This is especially true for multidimensional distributions.

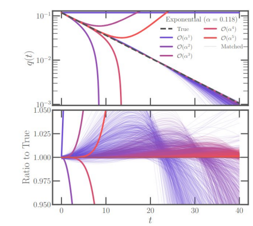


$$q(t_\alpha, t_\beta|\alpha_s) = \frac{\alpha_s C_F}{\pi(\alpha-\beta)} e^{\frac{\alpha_s C_F}{\pi(\alpha-\beta)}(t_\alpha - t_\beta) + \mathcal{O}(\alpha_s^2)} \times e^{-\alpha_s \frac{C_F}{2\pi}\left(\frac{t_\beta^2}{\beta} + \frac{(t_\alpha - t_\beta)^2}{\alpha-\beta}\right) + \int dt'_\alpha e^{\mathcal{O}(\alpha^2)}} \Theta(t_\beta < t_\alpha < \frac{\alpha}{\beta} t_\beta)$$

2. Constraining higher order terms due to unitarity
   a. Higher order terms must conspire to be normalized, finite, and positive.
   b. e.g. the alpha series can *not* truncate to zero

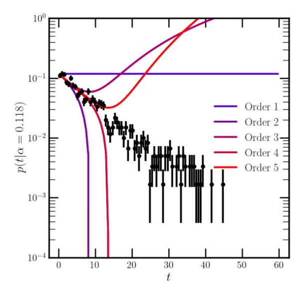
e.g. If the FO is $p_{\text{FO}}^{(M)}(t|\alpha) = \alpha \sum_m^{M-1} \frac{(-\alpha t)^m}{m!}$

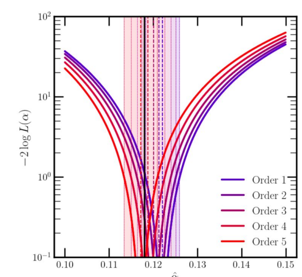
The full dist *must* have the form



$$q^{(M)}(t|\alpha) = \alpha \exp\left[-\sum_{m=M} g_m(t)\alpha^m - \alpha \int_0^t dt' \exp\left(-\sum_{m=M} g_m(t')\alpha^m\right)\right]$$

3. Non-pathological likelihood -> can do MLE and statistical fits to data *directly*, with no regularizations required!
   a. Higher order params have a nice interpretation as theory uncertainties / nuisances.





Analytic Matching

$$p(x|\alpha) = \underbrace{p_0(x) + \alpha p_1(x|\alpha) + \alpha^2 p_2(x|\alpha) + ... + \alpha^M p_M(x|\alpha)}_{\text{Known Fixed-order calculation}} + \underbrace{\mathcal{O}\left(\alpha^{M+1}\right)}_{\text{Unknown}}$$

Given an expression for the p_m, which can be either closed-form or numeric as a function of x or t, we can constrain part of the g function to reproduce the exact fixed-order behavior.

$$g(t,\alpha) = -\log\left(p_{m*}(t)\alpha^{m^*}\right) + \sum_k \frac{1}{k}\left(\left[-\sum_{m>m^*}^{M} \frac{p_m(t)}{p_{m^*}(t)}\alpha^{m-m^*}\right]^k - \left[\sum_{m=m^*}^{M}\int_0^t dt'\, p_m(t')\alpha^m\right]^k\right) + \mathcal{O}\left(\alpha^{M+1}\right)$$

Star = Divide out the lowest order such that the leading coefficient is 1

Can then collect powers of α

The higher-order part of g then represents the higher-order uncertainty — choose to be 0, your favorite numbers, or randomly according to your favorite prior! All are valid distributions*.

Looks scary, but autodiff magic helps us out here!

PyTorch

*As long as g doesn't diverge to positive infinity = the highest order coeffs are negative

Analytic Matching Toy: Exponentials

Solid: Fixed Order (Pathological)
Bands: Distribution of the RDF for different random choices of g_{mn}

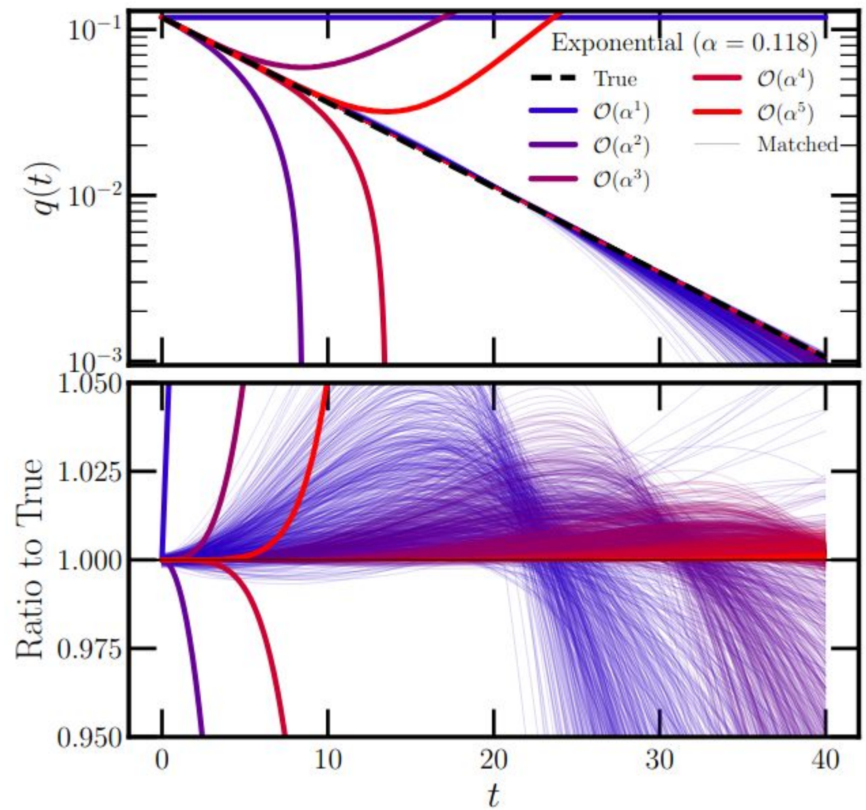
Consider the "platonic" distribution

$$p_{\text{Ideal}}(t|\alpha) = \alpha \exp(-\alpha t)$$

But we only have access to a finite fixed-order expansion

$$p_{\text{FO}}^{(M)}(t|\alpha) = \alpha \sum_{m}^{M-1} \frac{(-\alpha t)^m}{m!}$$

Let's apply the RDF! We'll pick some prior for the higher order coefficients *g*.



Full analytic answer: *All consistent distributions*

$$q^{(M)}(t|\alpha) = \alpha \exp\left[-\sum_{m=M} g_m(t)\alpha^m - \alpha \int_0^t dt' \exp\left(-\sum_{m=M} g_m(t')\alpha^m\right)\right]$$

Every curve in the bands is a *non-pathological* valid distribution that matches the FO!

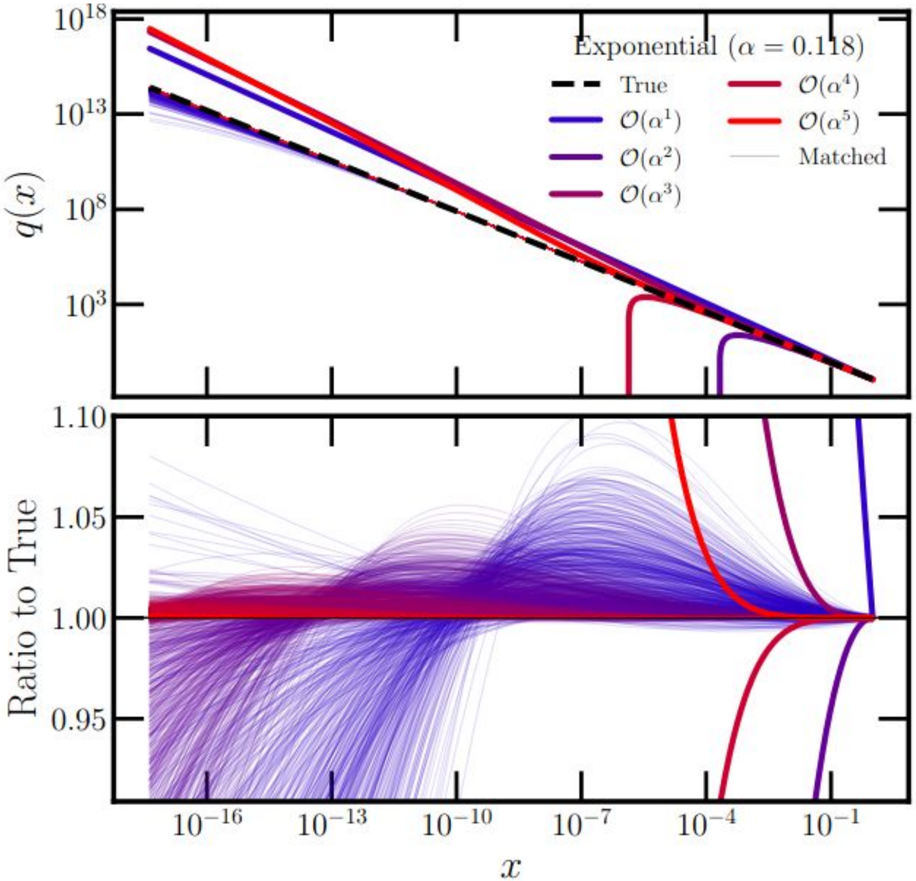
Analytic Matching Toy: Exponentials

An exponential distribution in *t* space is a power-law distribution in *x* space with a slope of α.

The F.O. distributions are only good for large *x*, and quickly blow up or become negative for small *x*.

But the RDF *never* will! It will always be well behaved.

Since the Ansatz is universal, the *true* answer corresponds to *some* choice of higher order *g*, but we don't know which — but no matter what we guess, our guess is always a valid PDF!



Analytic Matching Toy: Pseudo-Angularity

Consider the "platonic" distribution

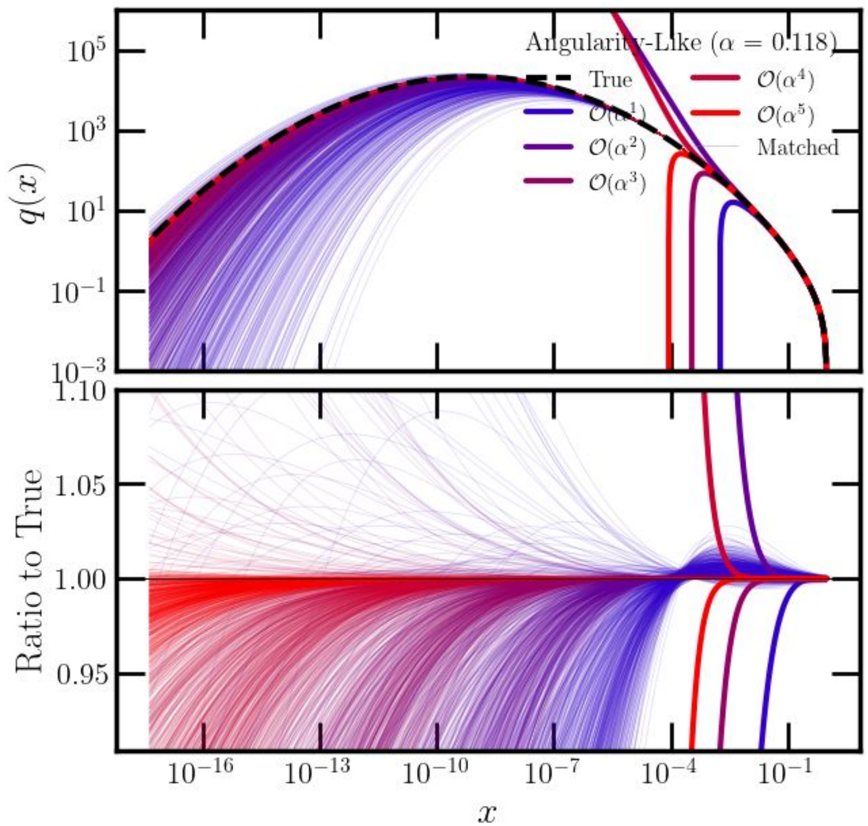
$$p^{\mathrm{LL}'}(\rho) = \frac{\alpha_s C_F}{\pi}\frac{1}{\rho}\left[\log\frac{1}{\rho}\right]\exp\left(-\frac{\alpha_s C_F}{2\pi}\left[\log^2\frac{1}{\rho}\right]\right)$$

E.g. Let's pretend that Leading-Log jet angularities are the way the world really works!

We can play the same game, only having access to fixed-order info.

We get *pretty good all-orders* predictions!

Solid: Fixed Order
Bands: Distribution of the RDF for different random choices of g_{mn}



Every curve in the bands is a *non-pathological* valid distribution that matches the FO!

Numeric Matching

The RDF still works even if we do not have access to a closed-form perturbative calculation! Instead, we can rely on numerical methods or event generators.

Given: Some method of constructing a *weighted histogram* of the observable for many different α, not necessarily normalized or even positive.

We pick a flexible-but-nice ansatz for *g* in terms of polynomials and single-logs*.

$$g(t,\alpha) = \log\left(\mathrm{ReLU}\left(\sum_{n=0}^{N-1} g_{0n} t^n \alpha^{m^*} \Theta(t-\theta_m)\right)\right) + \sum_{m=1,n=0}^{M-1,N-1} g_{mn} t^m \alpha^{m^*+m} \Theta(t-\theta_m)$$

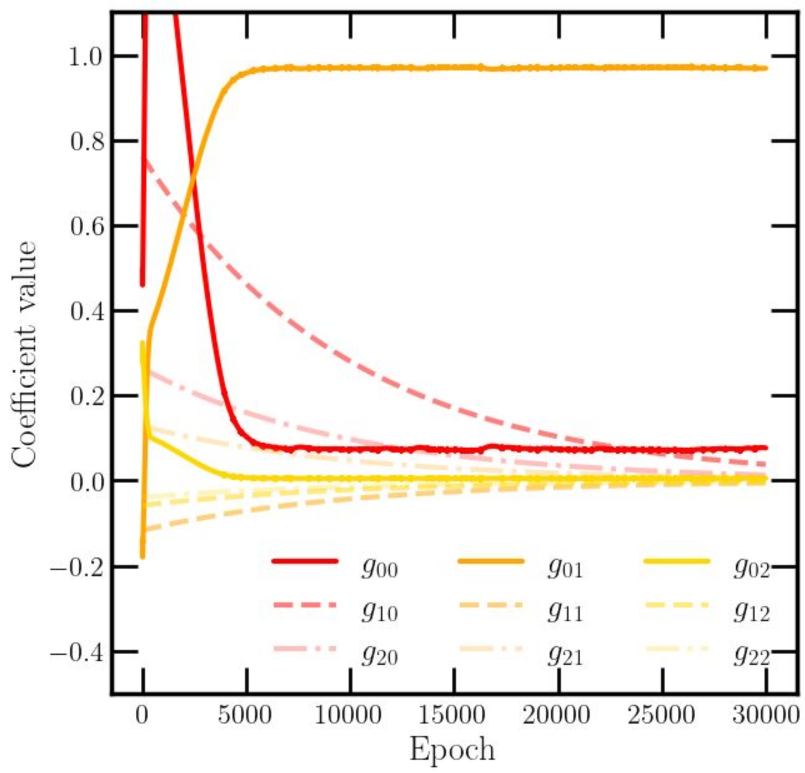
Then, we simply use the standard ML-bag-of-tricks** to minimize the MSE between (a) the fixed order histogram and (b) the Taylor expansion of the RDF for many α!

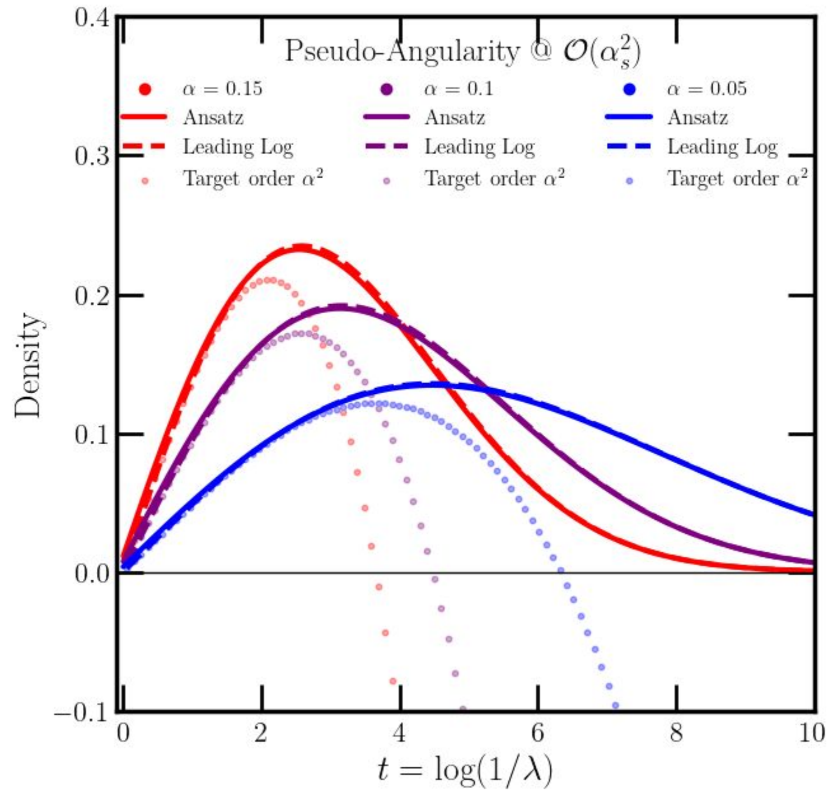
*In principle, *g* could have been anything. Even an entire neural network!

**See backup

Numeric Matching Toy: *Pseudo-Angularity*

Given a *weighted histogram* of events, we can perform approximate matching!

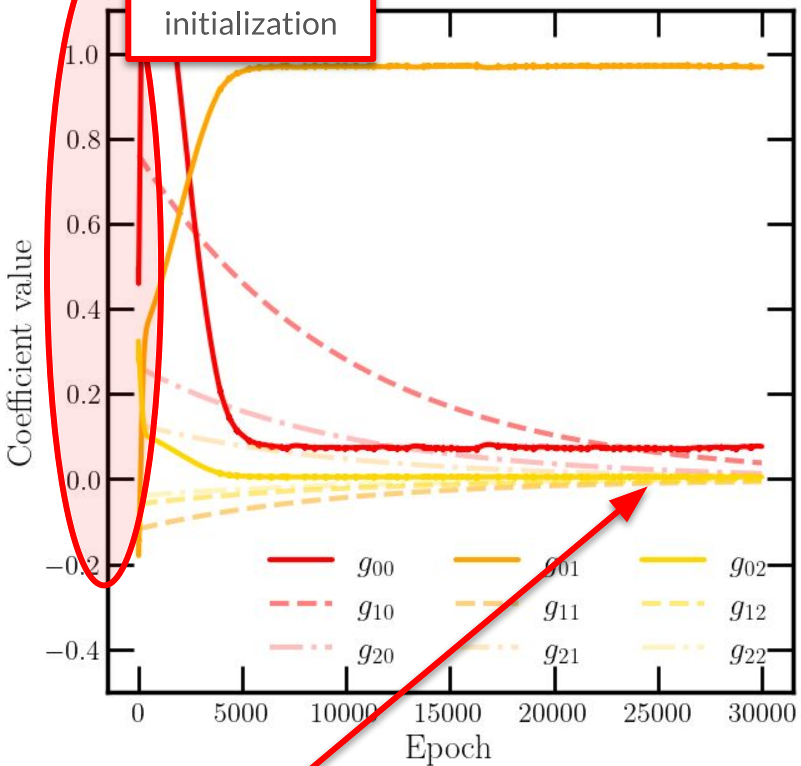




Using the F.O. expansion of $p^{\mathrm{LL}'}(\rho) = \frac{\alpha_s C_F}{\pi}\frac{1}{\rho}\left[\log\frac{1}{\rho}\right]\exp\left(-\frac{\alpha_s C_F}{2\pi}\left[\log^2\frac{1}{\rho}\right]\right)$

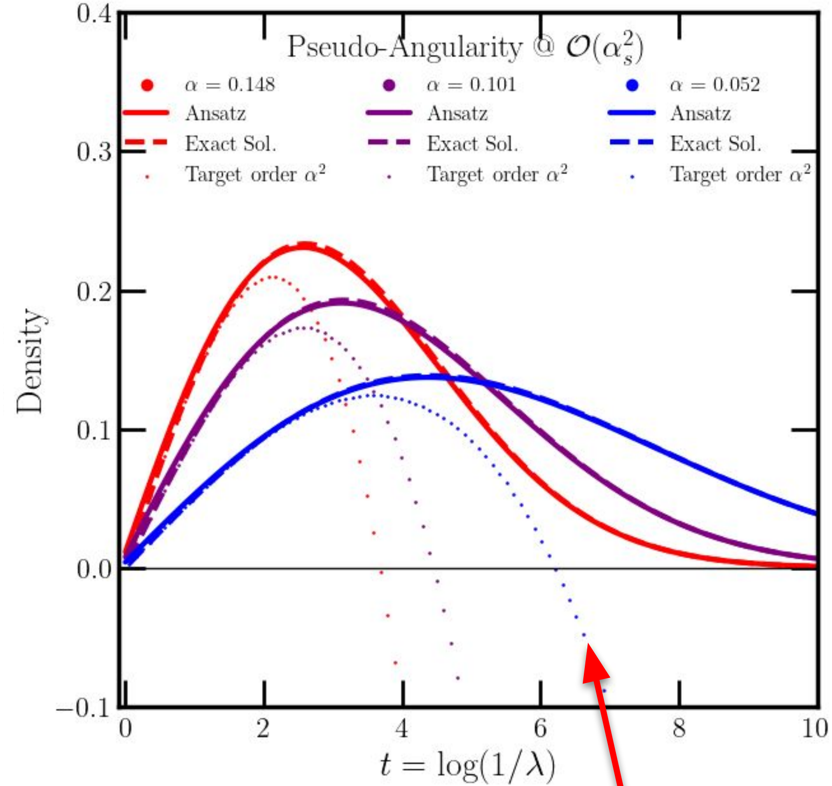
Numeric Matching Toy: *Pseudo-Angularity*

Given a *weighted histogram* of events, we can perform approximate matching!



Random initialization

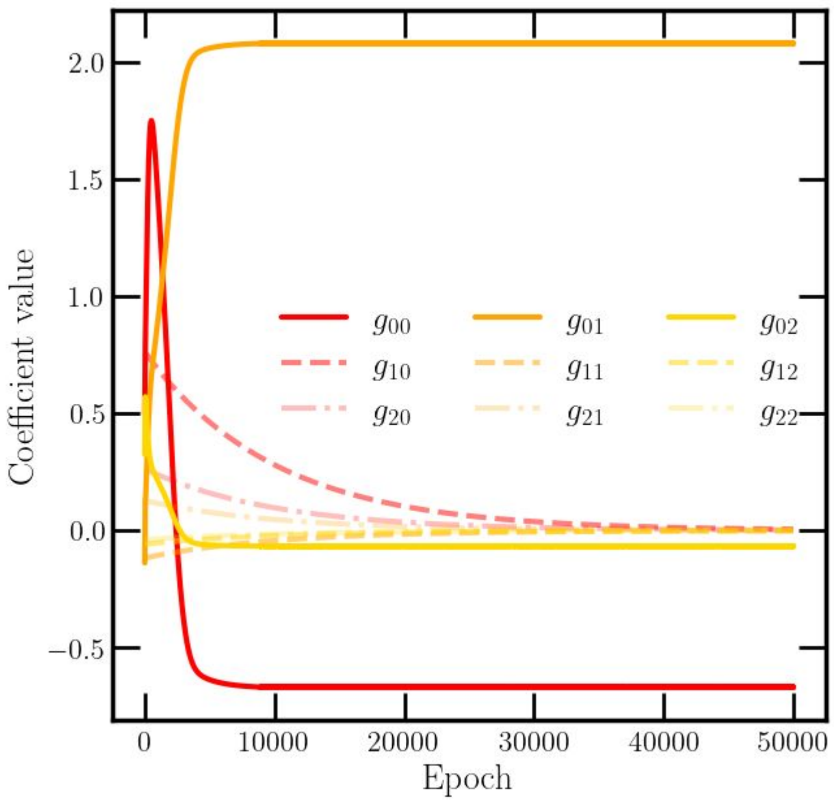
Coefficients of *g* functions learned using MSE fits such that the Taylor Series match.

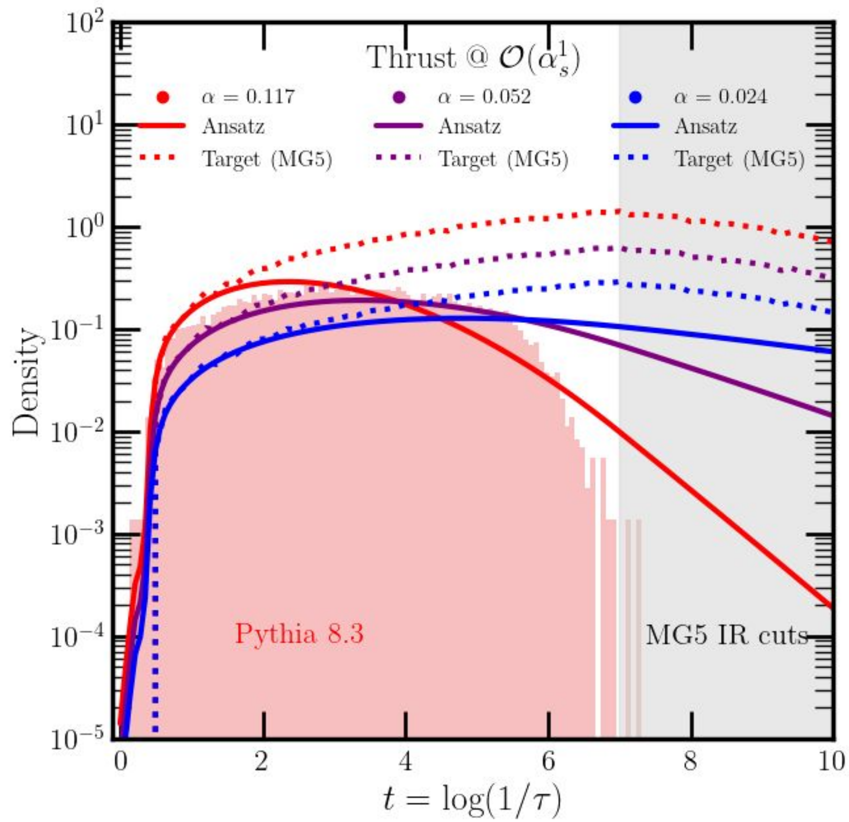


Weighted histogram of events — goes *negative* at 2nd order!

Numeric Matching: *Thrust*

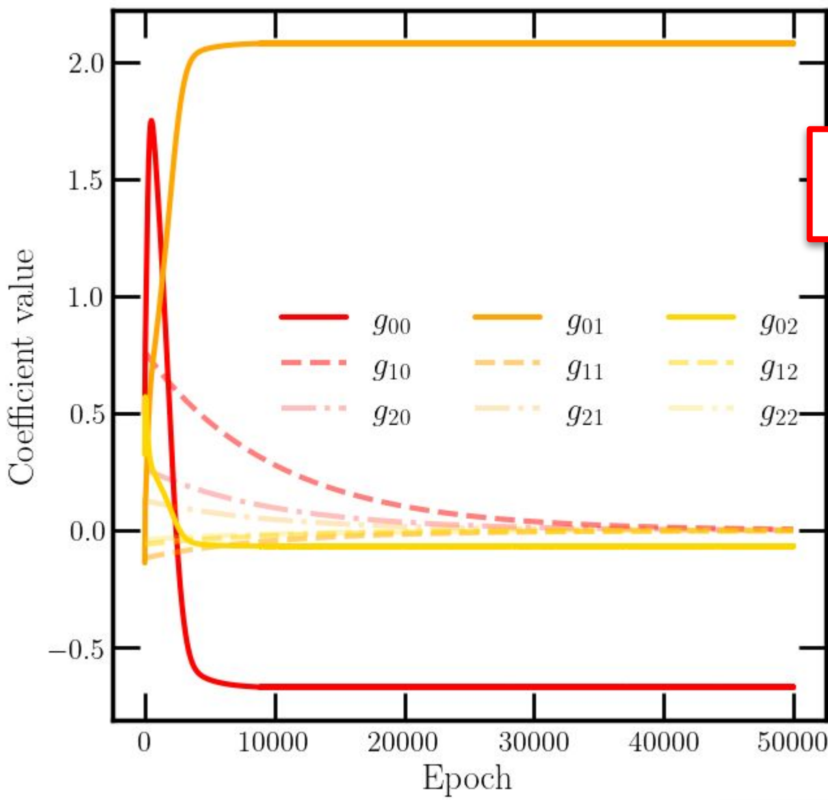
Realistic Events generated using a *Fixed-Order* generator (MG5), no showering

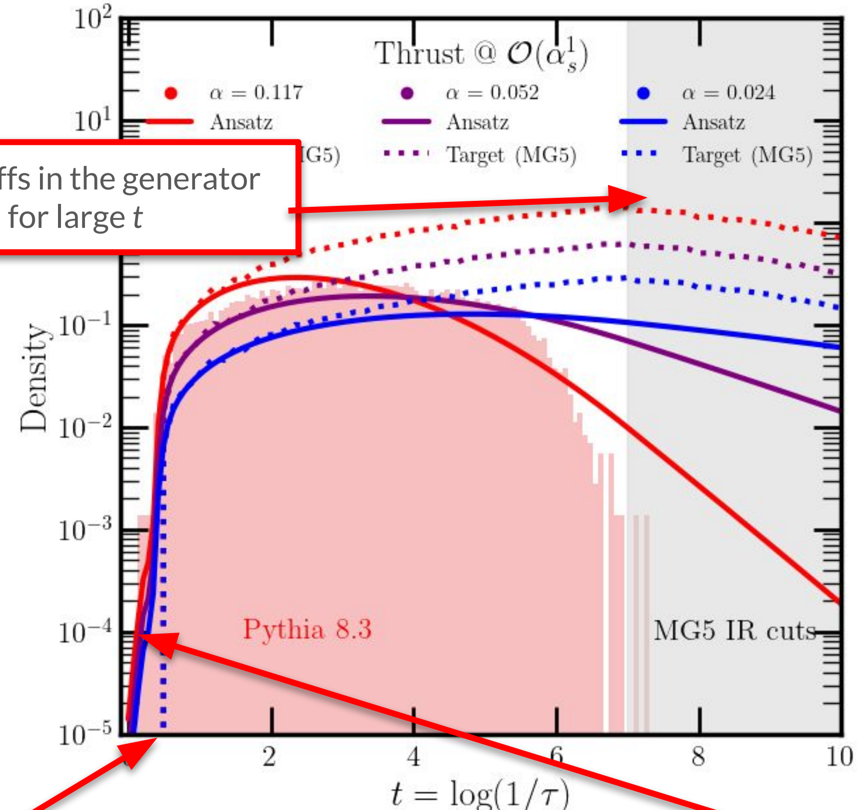




Numeric Matching: *Thrust*

Realistic Events generated using a *Fixed-Order* generator (MG5), no showering





IR Cutoffs in the generator for large *t*

Kinematic cutoffs — Thrust at first order does *not* saturate the full domain!

Doesn't match Pythia completely, but tracks in the high thrust region

Simultaneous Observables: Joint Angularity

Nothing special about single observables. The RDF works for multi-observables, in which resummation can be notoriously difficult!

Just split your joint F.O. calculation into an autoregressive product, and then do an independent RDF for each one, iteratively!

$$p(t_1, t_2, ..., t_k|\alpha) = p(t_1|\alpha) \cdot p(t_2|t1, \alpha) \cdot ... \cdot p(t_k|t_1, t2, ..., t_{k-1}, \alpha)$$

e.g. Joint Angularities

Soft-Coll Limit for Simplicity

$$p(t_\alpha, t_\beta) = \frac{\alpha_s C_F}{\pi(\alpha - \beta)}$$

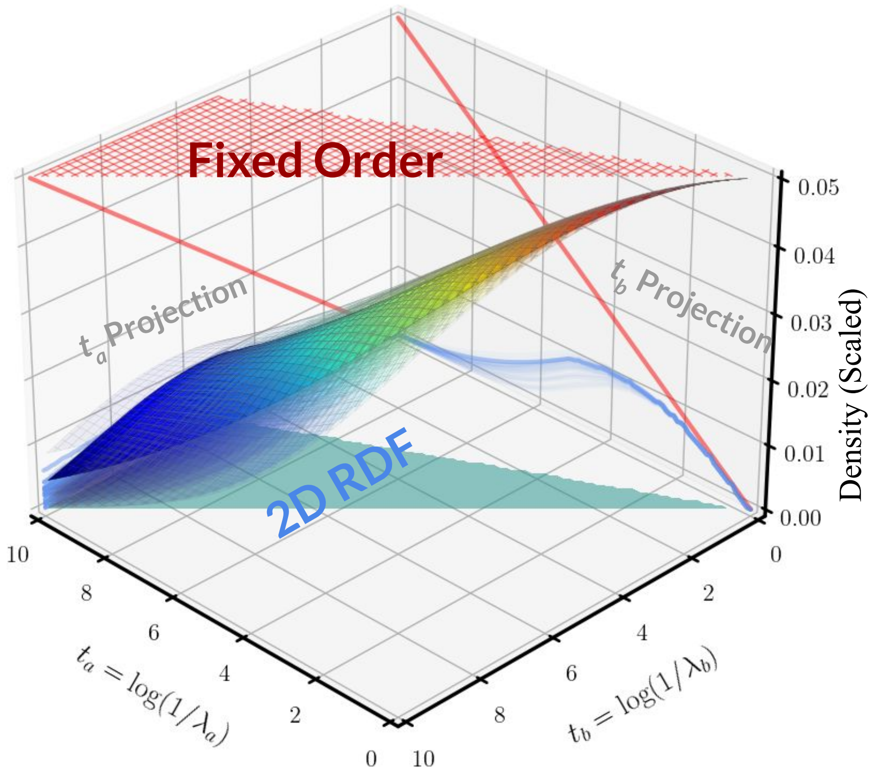
$$p(t_\beta) = \frac{\alpha_s C_F}{\pi \beta} t_\beta \Theta(1 - e^{-t_\beta})$$

$$p(t_\alpha|t_\beta) = \frac{\beta}{\alpha - \beta} \frac{1}{t_\beta} \Theta(t_\beta < t_\alpha < \frac{\alpha}{\beta} t$$

$$q(t_\alpha, t_\beta|\alpha_s) = \frac{\alpha_s C_F}{\pi(\alpha - \beta)} e^{\frac{\alpha_s C_F}{\pi(\alpha-\beta)}(t_\alpha - t_\beta) + \mathcal{O}(\alpha_s^2)}$$

Higher Order unconstrained terms

$$\times\, e^{-\alpha_s \frac{C_F}{2\pi}\left(\frac{t_\beta^2}{\beta} + \frac{(t_\alpha - t_\beta)^2}{\alpha - \beta}\right) + \int dt'_\alpha e^{\mathcal{O}(\alpha^2)}} \Theta(t_\beta < t_\alpha < \frac{\alpha}{\beta} t_\beta)$$



Matches the same double-logs as the full LL calculation!

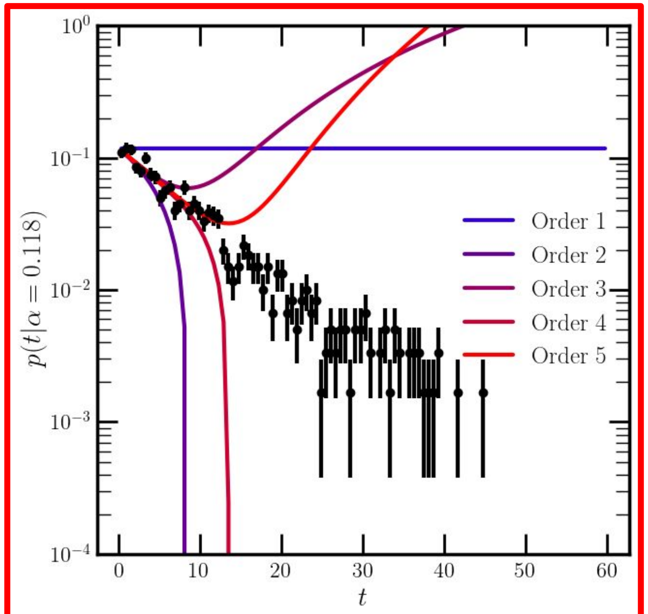
*Need to be extremely careful with Theta-Functions!

Inference: Fitting α From Data (Toy)

Toy: $p_{\text{Ideal}}(t|\alpha) = \alpha \exp(-\alpha t)$

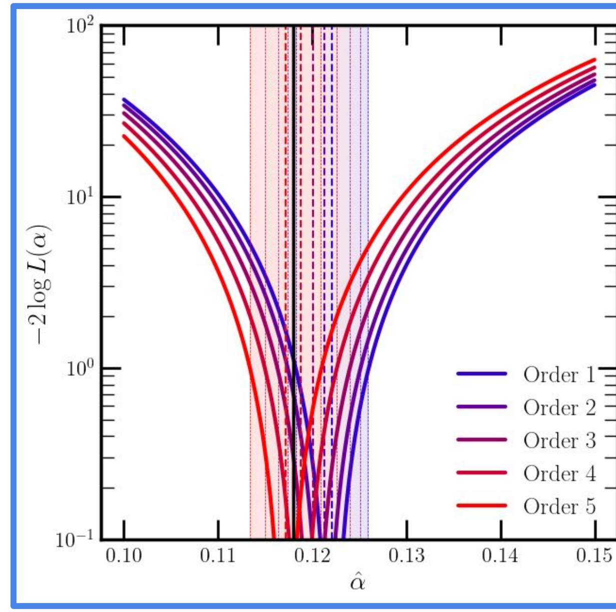
Goal: Fit α using only the F.O. information!

Since q are genuine, non-pathological PDFs, it's easy to *directly* do maximum-likelihood (or similar) inference!

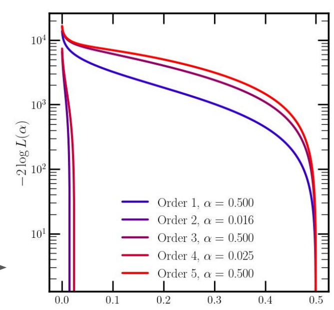


Fit with $q(x|\alpha)$ directly
(Full likelihood)

Higher order g's
=
Gaussian Nuisance Params



Direct F.O. likelihoods are so pathological that MLE is nonsensical



Cutoffs, binning, and other regulators needed!

CLAIMER

What we are claiming

- **Claiming:** Given a distribution p, pathological or otherwise, we can find a regulated distribution q that is normalized, smooth, positive, and normalized that is perturbatively matched to p.
- **Claiming:** This is a universal approximator: *any* valid q can be reached with a choice of higher order g, including the "true" one.
- **Claiming:** For *particular* g's, we can get meaningful q's that look like theory calculations and tame large logs!

DISCLAIMER

What we are *not* claiming

- **Not Claiming:** The perturbative structure is *easy* to maintain numerically or that numeric artifacts are manageable.
  - But if you restrict to reasonable choices of g, like $g \sim N(0,1)$, then the results seem reasonable!
- **Not Claiming:** The envelopes give a *prior-free* uncertainty on higher orders.
  - The envelopes only tell you what *possible* higher order physics is consistent with positivity and normalizability — it *constrains* physics
- **Not Claiming:** This beats logarithmic ressummation.
  - This is *an* all-orders summation. A genuine logarithmic resummation is always better — but for our method, you don't need to know *any* physics!

Key Point: This is *not* a shortcut to doing QCD, just a new way to parameterize things to guarantee nice properties. There is no free lunch, you still have to do physics!

Email us questions!
rikab@mit.edu
rmastand@berkeley.edu
Paper coming out on arXiv in August!

Conclusion

We introduced the Resummed Distribution Function (RDF) as a method to parameterize non-pathological distributions.

This ansatz is easy to match to fixed-order calculations and Monte Carlo, allowing one to (a) parameterize ignorance about higher order terms, (b) constrain and "guess" higher order terms via consistency conditions, (c) perform *an* all-orders resummation, if not *the* logarithmic resummation, and (d) build a non-pathological likelihood that can use used for fits

This method works analytically and numerically, on not just one observable but simultaneous observables in which resummation can be difficult.

This lets one peek at higher-order physics using nothing but consistency conditions, without needing to know anything about the structure of logs in QCD!

Appendices

Numerical Fitting

We generate events at first order in alpha in MG5_aMC for e+ e- > j j j at 500 GeV.

- Artificially reduce the jet p_T cutoff from 10 GeV to 0.25 GeV
- Generate ~200 different alpha values via reweighting
- No showering — this is a fixed order!

Fits are done to g by using the mean-squared-error between (a) the Taylor expansion of the RDF and (b) the F.O. histogram, batched over many values of alpha and t.

We implement this in PyTorch, using autodifferentiation and gradient descent (via ADAM) to minimize over the g_{mn} and Theta parameters.

Theta parameters are initialized to zero. g_{mn} are initialized with a normal distribution of mean zero and standard deviation 1/(m!n!).