JOINT INSTITUTE FOR NUCLEAR RESEARCH
Veksler and Baldin laboratory of High Energy Physics

FINAL REPORT ON THE INTEREST PROGRAMME

Puzzles of multiplicity

Supervisor:
Dr. Elena Kokoulina

Student:
Grecia Itzel Colín González, México
Michoacan University of Saint Nicholas of Hidalgo (UMSNH)

Participation period:
March 02 – April 19, Wave 14

Dubna, 2026

Table of Contents

Introduction **2**

Project Goals **3**

Scope of Work **4**

Methods **4**

Results **8**

Conclusion **10**

References **11**

Acknowledgments **11**

Introduction

High-energy particle collisions provide a unique window into the fundamental structure of matter and the dynamics of strong interactions. One of the most important observables in such processes is the multiplicity, defined as the number of secondary particles produced in a single event. Since particle production is inherently stochastic, multiplicity is described by a probability distribution P_n, normalized as

$$\sum_{n=0}^{\infty} P_n = 1, \qquad \langle n \rangle = \sum_{n=0}^{\infty} n P_n.$$

Rather than appearing as isolated particles, these secondaries are typically organized into collimated structures known as jets, which originate from the fragmentation of high-energy quarks and gluons. A jet can therefore be interpreted as the macroscopic manifestation of an underlying parton undergoing a cascade of successive emissions.

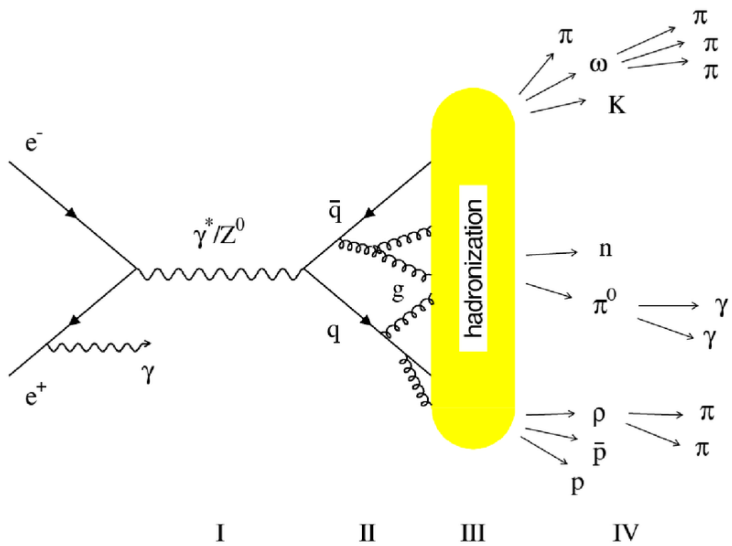

Figure 1: Schematic representation of e^+e^- annihilation: initial electroweak production, parton cascade, and hadronization stage.

The formation of jets involves two fundamentally different regimes. At high energy scales, the evolution of partons can be treated within perturbative Quantum Chromodynamics (pQCD), where the strong coupling constant depends on the energy scale,

$$\alpha_s(Q^2) \sim \frac{1}{\ln(Q^2/\Lambda^2)}.$$

In this regime, partons undergo successive splittings, forming a cascade in which energy is redistributed among an increasing number of degrees of freedom.

As the virtuality decreases, the system enters a non-perturbative regime where confinement becomes dominant. In this stage, partons cannot exist as free particles and instead transform into hadrons through the process of hadronization. This transition cannot be derived from first principles and is modeled phenomenologically.

A key feature of multiparticle production is its probabilistic nature. The cascade can be interpreted as a stochastic branching process, where each parton evolves independently. This perspective naturally leads to the introduction of the probability generating function,

$$Q(z) = \sum_{n=0}^{\infty} P_n z^n,$$

which encodes the full multiplicity distribution. Observable quantities such as the mean multiplicity and higher moments can be obtained through differentiation,

$$\langle n \rangle = \left. \frac{dQ}{dz} \right|_{z=1}, \quad \langle n(n-1) \rangle = \left. \frac{d^2Q}{dz^2} \right|_{z=1}.$$

Theoretical descriptions of jet evolution often rely on a two-stage picture. In the first stage, the parton cascade is described using perturbative QCD and stochastic methods. In the second stage, hadronization is introduced via phenomenological models. This framework provides a bridge between microscopic QCD dynamics and experimentally observed multiplicity distributions.

Project Goals

The objective of this work is to construct a complete theoretical framework describing multiparticle production by combining QCD-inspired branching processes with probabilistic methods. The goal is to derive the governing equations of the cascade, obtain multiplicity distributions, and incorporate hadronization in a consistent manner.

Scope of Work

This work focuses on the analytical description of electron–positron annihilation processes. The study includes the formulation of the cascade as a Markov process, the derivation of generating functions, and the implementation of a two-stage model connecting partons to hadrons.

Methods

The theoretical description of multiparticle production is based on the interpretation of the parton cascade as a stochastic Markov branching process. In this framework, the evolution of a jet is described as a sequence of independent parton splittings, where each parton evolves according to fixed probabilities that depend only on the current state of the system.

The development of the cascade is characterized by an evolution parameter Y, which can be interpreted as the effective thickness of the jet. It is defined as

$$Y = \frac{1}{2\pi b} \ln \left(1 + \alpha b \ln \frac{Q^2}{\mu^2} \right),$$

where Q^2 represents the virtuality scale and b depends on the number of colors and active flavors. As Y increases, the cascade evolves toward higher multiplicities through successive branching processes.

Within an infinitesimal interval ΔY, the cascade is governed by three fundamental QCD processes: gluon fission $g \to g + g$, quark bremsstrahlung $q \to q + g$, and quark pair creation $g \to q + \bar{q}$. These occur with probabilities $A\Delta Y$, $\tilde{A}\Delta Y$, and $B\Delta Y$, respectively. The assumption that each parton evolves independently and that these probabilities are constant defines the Markov property of the process.

Let $P_{m_g,m_q;n_g,n_q}(Y)$ denote the probability that a system initially containing m_g gluons and m_q quarks evolves into a state with n_g gluons and n_q

quarks after a thickness Y. The normalization condition requires that

$$\sum_{n_g,n_q} P_{m_g,m_q;n_g,n_q}(Y) = 1.$$

To encode the full statistical information of the cascade, generating functions are introduced. For a gluon-initiated jet:

$$G(u_g,u_q;Y) = \sum_{n_g,n_q} P_{1,0;n_g,n_q}(Y)\, u_g^{n_g} u_q^{n_q},$$

and for a quark-initiated jet:

$$Q(u_g,u_q;Y) = \sum_{n_g,n_q} P_{0,1;n_g,n_q}(Y)\, u_g^{n_g} u_q^{n_q}.$$

These generating functions provide direct access to multiplicity moments through differentiation. In particular,

$$\langle n \rangle = \left.\frac{\partial Q(z)}{\partial z}\right|_{z=1}, \quad \langle n(n-1) \rangle = \left.\frac{\partial^2 Q(z)}{\partial z^2}\right|_{z=1}.$$

The Markov nature of the process implies that the evolution satisfies the Chapman-Kolmogorov equation, which expresses the composition of probabilities over successive intervals. In terms of generating functions, this leads to functional relations of the form

$$G(Y+Y') = G(G(Y'),Q(Y');Y).$$

To derive differential equations governing the evolution, one considers an infinitesimal step ΔY. The generating functions are expanded as

$$G(u_g,u_q;Y+\Delta Y) = G(u_g + w^{(g)}\Delta Y,\ u_q + w^{(q)}\Delta Y;Y),$$

where $w^{(g)}$ and $w^{(q)}$ represent the infinitesimal generating functions corresponding to gluon and quark branching. These are constructed from the elementary processes and take the form

$$w^{(g)} = -(A+B)u_g + Au_g^2 + Bu_q^2,$$

$$w^{(q)} = -\tilde{A}u_q + \tilde{A}u_q u_g.$$

Expanding to first order in ΔY and taking the limit $\Delta Y \to 0$, one obtains the forward Kolmogorov equations:

$$\frac{\partial G}{\partial Y} = \frac{\partial G}{\partial u_g} w^{(g)} + \frac{\partial G}{\partial u_q} w^{(q)},$$

$$\frac{\partial Q}{\partial Y} = \frac{\partial Q}{\partial u_g} w^{(g)} + \frac{\partial Q}{\partial u_q} w^{(q)}.$$

These equations describe how the full probability distribution evolves as the cascade develops. An alternative formulation is obtained by considering the evolution of the generating function itself, leading to the backward Kolmogorov equations:

$$\frac{\partial G}{\partial Y} = -AG + AG^2 - BG + BQ^2,$$

$$\frac{\partial Q}{\partial Y} = -\tilde{A}Q + \tilde{A}QG.$$

These nonlinear equations form a closed system that completely determines the cascade dynamics once the initial conditions $G(u_g, u_q; 0) = u_g$ and $Q(u_g, u_q; 0) = u_q$ are specified.

The structure of these equations reveals that the cascade is intrinsically correlated. Unlike a simple Poisson process, where emissions are independent, the branching mechanism introduces correlations between produced particles. This leads to multiplicity distributions that are broader than Poisson and exhibit characteristic scaling properties.

In the special case where quark pair production is neglected ($B = 0$), the gluon cascade reduces to a pure branching process. The solution of the corresponding equation yields the generating function

$$G(u) = \frac{ue^{-AY}}{1 - u(1 - e^{-AY})},$$

which corresponds to the Furry-Yule distribution. From this, the mean multiplicity and dispersion are obtained as

$$\langle n \rangle = e^{AY}, \qquad D^2 = e^{AY}(e^{AY} - 1).$$

More generally, the multiplicity distribution of partons is well described by the negative binomial distribution (NBD),

$$P_m^{(p)} = \frac{\Gamma(m + k_p)}{\Gamma(k_p) m!} \left(\frac{\bar{m}}{\bar{m} + k_p} \right)^m \left(\frac{k_p}{\bar{m} + k_p} \right)^{k_p}.$$

This distribution emerges naturally from the stochastic cascade and incorporates correlations through the parameter k_p, which controls the width of the distribution.

The generating function associated with the NBD is

$$Q^{(p)}(z) = \left[1 + \frac{\bar{m}}{k_p}(1 - z)\right]^{-k_p}.$$

The second stage of the process corresponds to hadronization, which is modeled as a binomial process. In this picture, each parton produces hadrons independently, with a maximum number of hadron sources N and an average multiplicity $\bar{n}_h$. The probability distribution is given by

$$P_n^{(h)} = \binom{N}{n} \left(\frac{\bar{n}_h}{N}\right)^n \left(1 - \frac{\bar{n}_h}{N}\right)^{N-n}.$$

This form reflects the finite capacity of hadron production and ensures normalization. The corresponding generating function is

$$Q^{(h)}(z) = \left[1 + \frac{\bar{n}_h}{N}(z - 1)\right]^N.$$

For a system containing m partons, the independence of hadronization implies

$$Q_m^{(h)}(z) = \left[Q^{(h)}(z)\right]^m.$$

In electron–positron annihilation, the initial state includes two primary quarks, and gluons contribute differently to hadron production. This effect is incorporated through a parameter α, leading to an effective number of sources $(2 + \alpha m)$. Thus, the hadronization generating function becomes

$$\left[Q^{(h)}(z)\right]^{2+\alpha m} = \left[1 + \frac{\bar{n}_h}{N}(z - 1)\right]^{(2+\alpha m)N}.$$

Finally, the full multiplicity distribution is obtained by convolution of the cascade and hadronization stages:

$$P_n = \sum_{m=0}^{\infty} P_m^{(p)} \binom{(2 + \alpha m)N}{n} \left(\frac{\bar{n}_h}{N}\right)^n \left(1 - \frac{\bar{n}_h}{N}\right)^{(2+\alpha m)N-n}.$$

This expression represents the complete formulation of the Two-Stage Model. It encapsulates the interplay between the stochastic growth of partons in the cascade and the statistical nature of hadron formation, providing a consistent description of multiparticle production in high-energy processes.

Results

The numerical analysis was performed using a ROOT-based implementation of the Two-Stage Model, where the full multiplicity distribution P_n is computed through an explicit convolution of the parton cascade and hadronization stages. The model parameters are extracted by fitting the theoretical prediction to experimental multiplicity distributions using the Minuit2 minimization algorithm.

The parameters obtained from the fit correspond directly to those defined in the theoretical framework: the negative binomial parameter k_p, the average number of partons $\bar{m}$, the mean hadron multiplicity per source $\bar{n}_h$, the maximum number of hadrons N, the gluon contribution parameter α, and the normalization factor Ω. The fitted values for each center-of-mass energy are summarized in Table 1.

Table 1: Fitted parameters of the Two-Stage Model at different center-of-mass energies.

$\sqrt{s}$ (GeV)	k_p	$\bar{m}$	$\bar{n}_h$	N	α	Ω
14	14.0	0.08	3.87	17.7	0.97	2.00
22	3.15	2.41	4.60	26.5	0.19	2.00
34	6.96	3.58	5.89	12.5	0.95	2.00

The uncertainties of the fitted parameters were not explicitly included in this analysis. Although the minimization procedure internally computes parameter errors through the covariance matrix, these were not systematically extracted for all parameters in the implementation. Therefore, only central values are reported in order to avoid introducing inconsistent or non-reproducible uncertainties.

The experimental data used in the fits correspond to charged particle multiplicity distributions measured in e^+e^- annihilation experiments. In particular, the datasets implemented in the code are associated with measurements from the TASSO collaboration, as indicated in the graphical routines. These data provide the probability P_n of producing n charged particles at fixed energies.

The connection between the results and the theoretical framework described in the Methods section is direct. The numerical function implemented in the code evaluates the full expression of P_n, including the sum over the number of partons m and the binomial hadronization stage. In particular, the factor $(2 + \alpha m)N$ appearing in the computation represents the effective number of hadron sources, combining the two primary quarks with gluon contributions generated during the cascade.

For each energy, two complementary graphical representations of the multiplicity distribution are shown: a linear-scale plot and a logarithmic-scale plot. These plots allow a detailed comparison between the theoretical model and experimental data.

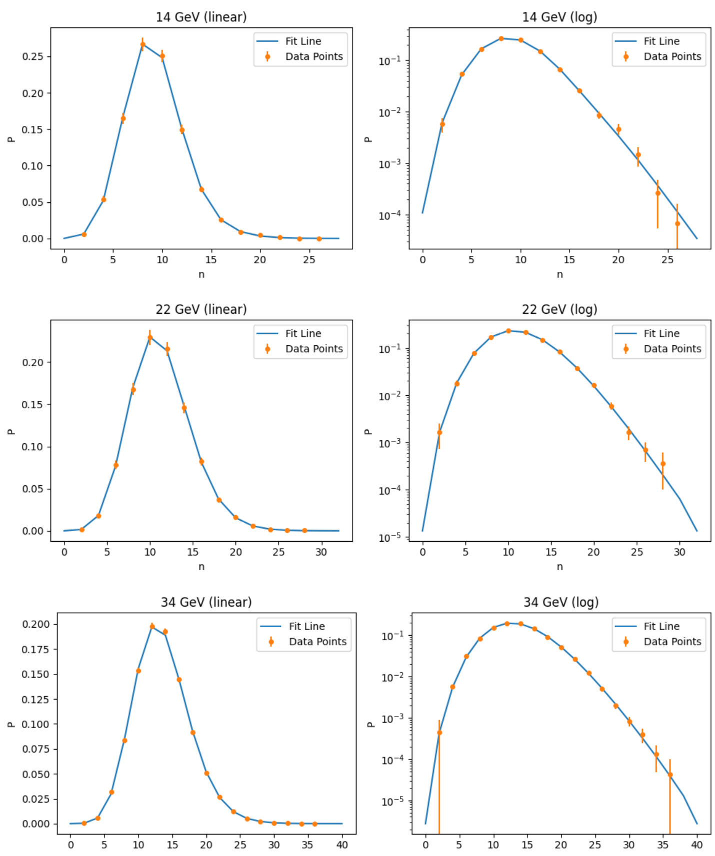

Figure 2: Multiplicity distributions P_n at $\sqrt{s} = 14$, 22, and 34 GeV. For each energy, the left panel shows the linear-scale distribution, while the right panel shows the same data in logarithmic scale. The points correspond to experimental data (TASSO), and the solid line represents the model prediction.

The linear-scale plots show the main structure of the multiplicity distributions. In all cases, the model successfully reproduces the position of the maximum and the overall shape of the distribution. As the center-of-mass energy increases, the peak shifts toward higher multiplicities and the distributions become broader, reflecting the increasing importance of parton branching processes.

The logarithmic-scale plots provide a more sensitive test of the model by emphasizing the tails of the distributions. In this region, corresponding to high multiplicities, the probabilities become very small and deviations from simple statistical models (such as Poisson distributions) become evident. The good agreement observed in the logarithmic plots indicates that the model correctly captures the correlations introduced during the parton cascade and the subsequent hadronization process.

A comparison with the code confirms that the shapes of the distributions are consistent with the implemented model. The smooth curves correspond to the theoretical function P_n computed through the convolution of the negative binomial parton distribution and the binomial hadronization stage, while the data points reproduce the experimental input arrays used in the fitting procedure. Minor differences in normalization or smoothness may arise depending on whether the plots are generated directly in ROOT or externally, but the physical behavior remains unchanged.

Overall, the results demonstrate that the Two-Stage Model provides a consistent and accurate description of multiplicity distributions in e^+e^- annihilation across different energies, capturing both the bulk behavior and the rare-event tails of the distributions.

Conclusion

The results obtained in this work show that the Two-Stage Model provides a consistent description of charged particle multiplicity distributions in e^+e^- annihilation at different energies. The model successfully reproduces both the main structure of the distributions and their high-multiplicity tails, indicating that the combined treatment of parton cascading and hadronization captures the essential physics of the process.

The evolution of the distributions with energy reflects the increasing role of parton branching, leading to broader multiplicity spectra. Despite the simplicity of the model assumptions, the agreement with experimental data supports its validity as an effective framework for describing multiparticle production.

Future improvements could include a more detailed extraction of parameter uncertainties and the study of correlations between parameters, allowing for a more complete statistical interpretation of the model.

References

[DELPHI Collaboration, 1996] DELPHI Collaboration (1996). Charged particle multiplicity distributions in restricted rapidity intervals in z0 hadronic decays. *Zeitschrift für Physik C.*

[Giovannini, 1986] Giovannini, A. (1986). *Multiplicity Distributions in High Energy Physics.* Springer.

[OPAL Collaboration, 2000] OPAL Collaboration (2000). Qcd studies with e+e- annihilation data at 172–189 gev. *European Physical Journal C.*

[TASSO Collaboration, 1983] TASSO Collaboration (1983). Jet production and fragmentation in e+e- annihilation at 12–43 gev. *Zeitschrift für Physik C.*

[TASSO Collaboration, 1989] TASSO Collaboration (1989). Charged multiplicity distributions and correlations in e+e- annihilation at petra energies. *Zeitschrift für Physik C.*

Acknowledgments

I would like to express my sincere gratitude to Dr. Elena S. Kokoulina for her support, patience, and guidance throughout this work. During the lectures and discussions, her clarity, kindness, and openness made the material accessible and created a welcoming learning environment. Even when I did not participate actively, her approach made it easier to follow the course, especially while adapting to both the language and the complexity of the subject.

I am also deeply thankful to the INTEREST program and the Joint Institute for Nuclear Research (JINR) for this opportunity. As my first experience in an international course, it was both intellectually stimulating and personally meaningful. It provided me with a new perspective and motivated me to continue exploring this field in future work.