From SU(3) to three quark families: Hard-learned lessons

To cite this article: Tommaso Dorigo 2020 Phys. Scr. 95 084004

View the article online for updates and enhancements.
From SU(3) to three quark families: Hard-learned lessons

Tommaso Dorigo

INFN, Sezione di Padova, Italy
E-mail: dorigo@pd.infn.it

Received 23 April 2020, revised 19 May 2020
Accepted for publication 15 June 2020
Published 23 June 2020

Abstract
The interplay between experimental breakthroughs and theoretical ideas that led to the discovery of hadron constituents and their classification into three families is discussed from a historical perspective, focusing on anecdotes that provide useful lessons from experimental physics practice.

Keywords: hadron structure, particle physics history, experimental design, standard model physics, data analysis

1. Introduction
This contribution aims at honoring the memory of Murray Gell-Mann, who recently passed away. Gell-Mann famously put forth in 1964 the idea of hadrons as composite objects made up of two or three constituents [1], which he called quarks. The scenery where this intuition came forth was a very confusing one, both theoretically and experimentally.

Theorists were at the time focusing on the paradigm of ‘nuclear democracy’, whereby baryons could be thought of as composite objects made up of mesons, and mesons themselves could be made of baryons. Little space was left to quarks in this picture, and the original idea of quarks was met with initial difﬁdence. Meanwhile, experimentalists were busy adding scores of entries to the list of known resonances through studies of fixed target reactions with more and more energetic and intense particle beams. Once Gell-Mann’s idea of a SU(3) group structure at the basis of hadron composition came about, their focus became that of discovering particles which would not ﬁt in that scheme. These were dubbed ‘exotic hadrons’.

I will start this short article by discussing three stories from that exciting, crazy period of the history of contemporary physics. The ﬁrst story comes from the search for free quarks, which was undertaken as a possible direct way to test Gell-Mann’s model. The second concerns the rise of the five-sigma prescription to which we nowadays diligently stick in determining whether an experimental observation of data not conforming to the accepted theoretical model is worth the label of discovery of new physics. The third is the very discovery of the charm quark, which gave a ﬁnal conﬁrmation to the quark model. I will then complete the discussion with a short recollection of the discoveries of the bottom and top quark, which completed the picture of hadronic constituents of matter.

2. Evidence of quarks in air-shower cores
The title of this section is the same as that of an article appeared on Physics Review Letters in 1969 [3]. In it, authors McCusker and Cairns argued to have discovered free quarks in cosmic ray showers detected by an airborne Wilson chamber. Already one year before, they had observed four charged tracks whose apparent ionization was compatible with the one expected from particles of charge equal to two-thirds of the electron charge. In their 1969 paper they showed a picture of droplets left by a particle which could arguably be nothing else but a fractionary charge particle: the track was made up by 110 droplets against an expectation of 229. The latter was the average obtained from a set of 55,000 observed tracks.

We may easily perform a frequentist calculation of the p-value of observing as few or less than 110 droplets from a Poisson process that yields on average 229: we only need to...
From the above, the Bonferroni-corrected p-value is
\[ p(N \leq 110) = \sum_{i=0}^{110} \frac{229^i e^{-229}}{i!} \approx 1.6 \times 10^{-18} \]  
\[ p' = 1 - (1 - p)^{55000} \approx 10^{-13}. \]

A Bonferroni correction [4] accounting for the multiple testing of 55,000 tracks should be applied to the above number, obtaining for the probability that at least one track has 110 or fewer droplets the value \( p' \) given by
\[ p' = 1 - (1 - p)^{55000} \approx 10^{-13}. \]

While the above calculation is in principle correct, its basic assumption is deeply flawed. Indeed, a basic knowledge of Statistics is required to avoid the pitfall into which McCusker and Cairns fell. As a good nuclear physicist might well know, the formation of droplets and the scattering process of charged particles in a medium are independent phenomena; each can be described by a Poisson distribution function. Droplets formation may be modeled by assuming that on average each scattering interaction produces four droplets. The combination of two Poisson processes is not equivalent to a single one. A correct treatment of the p-value calculation should involve the integration of the compound Poisson distribution, such as
\[ p(N \leq 110) = \sum_{i=0}^{110} \sum_{N=0}^{\infty} \left( \frac{(Np)^i e^{-Np}}{i!} \right) \left( \frac{\lambda^N e^{-\lambda}}{N!} \right) \approx 4.7 \times 10^{-5} \]
\[ p' = 1 - (1 - p)^{55000} = 0.925. \]

From the above, the Bonferroni-corrected p-value is \( p' = 1 - (1 - p)^{55000} = 0.925 \). The conclusion is that the observation of a low ionization track is absolutely normal. This striking mistake is a clear real-life example of the dangers of misunderstanding the probability distribution function from which one’s data are sampled. As for free quarks, they were never found in air showers or other laboratory experiments, and the theory of Quantum Chromodynamics (QCD) developed in the seventies [5, 6] soon provided the fundamental reason why those attempts were destined to fail.

### 3. The birth of the five-sigma criterion

In 1968 Arthur Rosenfeld wrote an article titled ‘Are There Any Far-out Mesons or Baryons?’ [7]. In the jargon of particle physics in the sixties, ‘far-out hadrons’ indicated hypothetical hadrons not fitting in SU(3) multiplets—the above-mentioned exotic hadrons. In the article, Rosenfeld demonstrated that the number of claims of discovery of such exotic particles published in scientific magazines in the sixties agreed reasonably well with the number of statistical fluctuations that one could expect to observe in the analyzed datasets. He examined the literature and pointed his finger at the large ‘trials factors’ (the sizes of required Bonferroni corrections [4]) coming into play due to massive use of combinations of observed particles in the final state of collisions in the derivation of mass spectra containing potential discoveries:

‘[...] This reasoning on multiplicities, extended to all combinations of all outgoing particles and to all countries, leads to an estimate of 35 million mass combinations calculated per year. How many histograms are plotted from these 35 million combinations? A glance through the journals shows that a typical mass histogram has about 2,500 entries, so the number we were looking for [...] is then 15,000 histograms per year [...]’.

‘[...] In summary of all the discussion above, I conclude that each of our 15,000 annual histograms is capable of generating somewhere between 10 and 100 deceptive upward fluctuations [...]’.

Rosenfeld could thus conclude his line of reasoning with a suggestion which should be recognized for its foundational value for the five-sigma criterion now commonly used in the practice of hypothesis testing in high-energy physics (HEP) and related fields:

‘To the theorist or phenomenologist the moral is simple: wait for nearly 5σ effects. For the experimental group who has spent a year of their time and perhaps a million dollars, the problem is harder... go ahead and publish... but they should realize that any bump less than about 5σ calls for a repeat of the experiment.’

The rationale behind the original proposal of a five-sigma criterion for discovery claims was thus that of addressing the multiple testing practice in HEP. Yet more recently the criterion gradually came to be used to rather provide protection from unaccounted systematic uncertainties, as the discoveries of e.g. single top quark production at the Tevatron in 2009 [8, 9] or of the Higgs boson at the LHC in 2012 [10, 11] testify.

### 4. Lederman’s shoulder

In the mid-sixties it was observed that the 30 GeV protons of the Alternating Gradient Synchrotron copiously produced muons in beam-dump interactions; these were initially believed to come from W boson decays [12]. However, Y. Yamaguchi and L. Okun in 1966 pointed out that these muons could also be the result of virtual photon interactions [13]. This lent itself to be investigated in detail with a rather simple experimental setup, which was exploited in the following years by Lederman, Limon, Christenson, and Zavattini [14]. The team of physicists studied collisions of the 30 GeV protons with Uranium and sought for muon pairs, in order to understand the properties of virtual photon production of lepton pairs and to see if the expected smooth invariant mass spectrum of those particle pairs would be interrupted by vector meson resonances. Unfortunately, the spectrometer they employed was a rather crude device, which determined the momentum of muons through the range of those particles in thick layers of steel. The resulting mass resolution of the apparatus was thus insufficient to allow for a detailed investigation of the spectrum:

‘The yield of muon pairs decreased rapidly from 1 GeV to the kinematic limit of nearly 6 GeV with the exception of a curious shoulder near 3 GeV. The measurement of muons was
by range as determined by liquid and plastic scintillation counters interspersed with steel shielding. Each angular bin (there were 18) had four range bins, and for two muons this made a total of only 5000 mass bins into which to sort the data. Multiple scattering in the minimum of 10 feet of steel made finer binning useless. Thus we could only note that ‘Indeed, in the mass region near 3.5 GeV, the observed spectrum may be reproduced by a composite of a resonance and a steeper continuum.’ [12]

Clearly, the signal of a large number of $J/\psi \rightarrow \mu^+ \mu^-$ decays was there; yet unfortunately it could not be exploited to further our knowledge of fundamental matter. The lesson to be learned is clear: energy resolution is of paramount importance. Lederman made good use of it when he directed the effort that led to the discovery of the fifth quark, beauty, in 1977 [15]. Before that, however, the charm quark would make its second, decisive apparition at Brookhaven and SLAC, again as a bound quark-antiquark state in the guise of the $J/\psi$ meson.

5. The charm discovery

In the early seventies two independent theoretical ideas already suggested that quarks could be more than three. Maybe four, or maybe even at least six. These ideas both had an experimental input. An extension to four flavours of the eightfold way was suggested by Glashow, Iliopoulos and Maiani [16] to account for the failure to observe neutral kaon decays to muon pairs; and the small CP asymmetry violation in weak decays of neutral kaons observed already in 1964 by Christenson, Cronin, Fitch, and Turlay [17] was the input to a speculation of a three-family structure of the quark matrix by Kobayashi and Maskawa [18], who followed the work of Cabibbo extending his famous parametrization of allowed and suppressed decays through the angle that still bears his name [19]. I will briefly discuss below the basic ideas of the first of these theoretical breakthroughs, which paved the way to the discovery of charm and to the definitive consacration of Gell-Mann’s static quark model of hadrons.

In Gell-Mann’s model the charged kaon ($K^+$) is composed by an up quark and an anti-strange quark. Its mass of 494 MeV and it decays mostly into muon-neutrino pairs (63% of the time), or to two pions ($\pi^+ \pi^-$). The neutral kaon ($K_0$, e.g.) is instead made up by a down quark bound to an anti-strange quark. Its mass is of 498 MeV, and it commonly decays to various final states (three pions, or a pion accompanied by an electron-neutrino or muon-neutrino pair). The branching fraction of the neutral kaon to muon pairs is quite rare, at $7 \times 10^{-7}$. The question was then, if there exist both charged and neutral currents, as the recent electroweak model of Glashow, Weinberg and Salam implied [20–22], what prevents neutral kaons from decaying into muon pairs? To shed light into this matter one needs to recall the role of the Cabibbo angle in weak interaction rates.

If one studies neutron and Lambda baryon decays, one observes a similar structure in the static quark model: they supposedly proceed through an anti-up–down quark current or an anti-up–strange quark current, respectively. However, experimental tests indicate that the latter has an intensity which is twenty times smaller. This is annoying, as we would like to describe the interaction as a product of $V–A$ currents: in terms of matrix elements $M$ we could write

$$M_{d\rightarrow u} = \frac{g}{\sqrt{2}} \left[ \bar{\psi}_u \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_d \right]$$

$$\times \frac{1}{m^2 - q^2} \frac{g}{\sqrt{2}} \left[ \bar{\psi}_u \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_d \right]$$

and

$$M_{s\rightarrow u} = \frac{g}{\sqrt{2}} \left[ \bar{\psi}_u \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_s \right]$$

$$\times \frac{1}{m^2 - q^2} \frac{g}{\sqrt{2}} \left[ \bar{\psi}_u \gamma^\mu \frac{1}{2} (1 - \gamma^5) \psi_s \right]$$

where $m$ and $q$ are the mass and four-momentum of the propagator, and $\psi$ spinors describe fermion fields. In the above relations, the coupling constant $g$ associated to the two vertices is the same, and yet there is a factor of 20 in the relative intensity. The universality of the charged weak current can be retained, as suggested by Nicola Cabibbo in 1963 [19], if a mixing of strong interaction eigenstates ($u$, $d$, $s$) is operated by the angle $\theta$, producing the weak interaction eigenstate $d' = d \cos \theta + s \sin \theta$. This indicates that strong and weak interactions see different properties of quarks. Following this transformation, amplitudes of weak neutral currents with no change in strangeness ($\Delta S = 0$) or unit change of strangeness ($\Delta S = \pm 1$) depend respectively on $\cos \theta$ and $\sin \theta$, and intensities and lifetimes on the square of those factors. For the Cabibbo angle one could then estimate a value of about 13 degrees, as e.g. from the ratio of the relative widths of Lambda and neutron,

$$\Gamma(\Lambda \rightarrow p e \nu) / \Gamma(n \rightarrow p e \nu) \approx \frac{g^4 \sin^2 \theta}{g^4 \cos^2 \theta}$$

after accounting for phase space factors.

The mixing of $d$ and $s$ quarks implies the existence of neutral currents changing quark flavour, endowed with $\cos \theta$ and $\sin \theta$ factors. This comes in from writing a neutral current as

$$J^0 = \psi_u \bar{\psi}_u - \psi_d \bar{\psi}_d$$

and explicating the terms $d' = d \cos \theta + s \sin \theta$. The above neutral current should enable the decay to two muons of the neutral kaon, which experiments however show to be heavily suppressed, as mentioned above. In 1970 Glashow, Iliopoulos and Maiani solved this problem radically, by postulating what is now dubbed the ‘GIM mechanism’: the existence of a fourth quark may cancel the mixing contribution, making neutral currents incapable of changing quark flavours. An incomplete cancellation in box diagrams can be attributed to different quark masses of up and charm; from this and observed rates of neutral kaon decays, the theorists came up with the prediction that the fourth quark had to have a mass in the 1 to 3 GeV range.

Despite the existence of machines capable of producing the $J/\psi$ particle and of detector technologies largely sufficient
for a precise interpretation of its decays, the discovery of the charm quark, and the subsequent acceptance of quarks as the fundamental constituents of hadronic matter, had to wait until 1974, when reports were given of a resonance at 3.1 GeV in the invariant mass of muon pairs produced in proton collisions on a Beryllium target [23] and a huge enhancement in the annihilation cross section of electron-positron pairs at the corresponding energy [24]. The latter effect was also immediately confirmed by the ADONE accelerator in Frascati [25]. It was very unfortunate for the Italian facility to have opted to not push to their highest achievable energy the beams before then: it will be recorded as one of the most striking cases of a safety-based decision on an experimental research plan turned out to have very significant, unwanted consequences.

6. On to the bottom quark

The charm discovery was soon followed by the observation of the $\tau$ lepton [26] by Martin Perl’s team, again in electron-positron collisions at Stanford. The observation of the $e^+e^- \rightarrow \tau^+\tau^-$ process remained controversial for some time: due to the presence of multiple neutrinos in the final state the produced events lacked the smoking-gun signature of an invariant mass peak. Regardless of the still undecided nature of the $\tau$ signal, by the mid-1970s a consensus started to form around the idea that a third family of quarks was there to be found. The biggest indicium was clearly the CP violation effect observed over a decade earlier: that observation could only fit in the Standard Model if a complex phase could find place in the quark mixing matrix, something which only three or more generation of matter fields allowed. Several experiments started looking for a bound state of a new, still heavier quark and its antiparticle: this was correctly understood to be the cleanest way to proceed. The best positioned for such a quest was the E288 experiment at the Fermi National Accelerator Laboratory, where Leon Lederman’s team constructed a double-arm muon spectrometer suitable for collisions provided by the 400 GeV protons of the beam produced by the main ring, at the time the highest-energy proton synchrotron in operation.

Concerning the E288 search, it is interesting to recall that in 1976 the group published an article where they claimed to have found a resonance with a mass of 6 GeV [27]. The accumulation of reconstructed muon pairs at that mass value was not really very compelling: in fact, in the article they noted that they studied with toy experiments the frequency with which such an effect could be observed from spurious fluctuations:

‘Clusters of events as observed occurring anywhere from 5.5 to 10.0 GeV appeared less than 2% of the time. Thus the statistical case for a narrow (<100 MeV) resonance is strong although we are aware of the need for a confirmation.’

The true Upsilon discovery followed quite un emulation thereafter, one year later [15]: cautioned by their previous incorrect claim, the E288 scientists waited for a long time after seeing a very clear, towering peak at 9.5 GeV. They performed a large number of statistical tests, estimating the significance of their Upsilon signal at 3, 5, then 8 standard deviations. Eventually, by the time they decided to publish their findings, they did not even offer an estimate of significance: indeed, the signal easily passed what experimental physicist sometimes call the ‘inter-ocular stress test’: if it hits you between the eyes, you know it is there. The signal was later well resolved into its three peaking components, later confirmed to be due to muon-pair decays of the 1S, 2S, 3S states of the bottomonium vector meson. The properties of the new quark were successively studied also at electron-positron facilities (PLUTO/DASP), confirming charge and weak isospin assignments of the new quark [28].

7. The race to the top

Once a bottom quark was established, there were no doubts that a sixth quark, top, was also needed. There are several reasons for this:

• the renormalizability of the theory demands existence of Ward identities, which in turn require cancellation of anomalies from triangle diagrams. These diagrams connect two vector and one axial-vector current through fermion loops; since all fermions contribute to the total coupling, the cancellation does not take place if the bottom quark is a weak isosinglet [29].

• In analogy to the GIM mechanism, a large branching fraction (12%) of neutral B mesons into dilepton pairs can be predicted if the bottom quark is an isosinglet; this was ruled out by the UA1 experiment at CERN and by the CLEO experiment at Cornell, which showed that there were no detectable dileptonic decays of neutral B mesons [30].

• A number of measurements of electroweak observables performed at electron-positron machines at $\sqrt{s}$ values of tens of GeV, in particular forward-backward asymmetries, later provided indirect estimates of the weak isospin of the bottom quark, which were only compatible with the $t^0_{\text{fl}} = -0.5$ assignment required by a three-family model [31].

• In addition to the above, the ARGUS collaboration measured a large mixing between neutral B mesons [32]. This indicated that a sixth quark had to contribute to the amplitude, and that its mass had to exceed the value of approximately 40 GeV.

At the turn of the decade Fermilab responded to the conversion of CERN’s Super-proton-synchrotron to a proton-antiproton collider with plans for an even higher-energy machine. After considering a number of possible design choices, the laboratory director Robert Wilson settled on the idea of constructing a 2-TeV proton-antiproton collider, and put the detector project in the hands of Alvin Tollestrup. Alvin understood the need to create an international collaboration around it. So he convinced Japanese physicists to join, famously asking Kunitaka Kondo to provide ‘five people and five million dollars’ [33]; the Japanese team would later
grow to constitute a steady 15% of the total manpower of the CDF experiment for the whole duration of that longevage experiment. In 1980 Giorgio Bellettini also joined the construction effort with a few Italian colleagues. The experiment gradually grew to several hundred collaborators, with a quite significant Italian component and members from all over the world.

The CDF detector was built in the early eighties, and started operations in 1985, but only collected a significant dataset of $\sqrt{s} = 1.8$ TeV proton-antiproton collisions in its 1987-88 ‘run 0’. By that time the UA1 collaboration at CERN produced and soon retracted a top-quark-discovery claim [34], based on the observation of a dozen events containing a W boson decay with hadronic jets. The modeling of QCD radiation was still in its infancy back then, leading UA1 scientists to mistake those events as decays of W bosons to a 40GeV top quark. A similar pitfall occurred in 1992, when a clean dilepton event candidate observed in run 0, and consistent with the decay chain $t\bar{t} \rightarrow e^+\nu_e b\bar{u} \rightarrow \nu_e \bar{u} b\bar{u}$, was the basis of a kinematic analysis by a team independent from CDF, which supported that interpretation [35]. In retrospect that event is almost certainly to be attributed to background sources; yet that kinematic analysis method, which got around the unconstrained nature of the dilepton final state of top pair decay by a reweighting technique, outlived the spurious claim and remains to this day a valid reconstruction tool.

Run 1 of the Tevatron, which started in the summer of 1992, saw CDF joined by its competitor across the ring, DZERO. CDF could now deploy the first silicon microvertex detector ever built to operate in the harsh environment of proton-antiproton collisions, the SVX; this offered a large advantage in the top hunt, as tracks originated from long-lived B hadrons (created by the hadronization of bottom quarks resulting from the $t \rightarrow Wb$ decay) could be identified by their large impact parameter. DZERO had also another big shortcoming: it had been designed with the primary purpose of precisely measuring the $W \rightarrow e\nu$ decay, and for that purpose it featured a very performant uranium-liquid argon calorimeter which allowed precise estimates of the energy of electrons; the design however prevented the endowment of the central detector with a magnetic field. DZERO could therefore not measure the momentum of charged particles (except muons through dedicated outer detectors), which was a significant disadvantage with respect to CDF.

By mid-1993 an Italian group of CDF collaborators led by Giorgio Bellettini claimed that they were seeing clear evidence of top quark pair decays in the kinematics of jets produced in association with a leptonic W signal, in the 20 inverse picobarns of data until then collected [36]. A new controversy followed, as the experiment leaders sabotaged the analysis by delaying its approval via endless requests of new checks and verifications. Their untold stand was that CDF had to discover the top by showing the signal of bottom quarks produced by top decays, which the SVX enabled. The information from the kinematic analysis was not allowed to become public, although the events singled out by that search were basically the same ones selected by the mainstream analysis, which utilized b-tagging criteria [37].

By early 1994 CDF could isolate selected event samples where globally an excess of about three standard deviations could be estimated over background sources. That evidence was compounded by a similarly sized evidence coming from the mass reconstruction enabled by $W + 4$-jet events. However, at that critical juncture, the over-conservative mindset of the CDF collaboration kicked in, and a paper was published in April 1994 with the word ‘evidence’ in the title [33, 38]: for the first time, Rosenfeld’s 5-sigma criterion went into effect, and CDF fell short of declaring victory in the top hunt. Still, it is remarkable to point out how that early analysis (based on just a handful of events, only 7 of which enabled a mass measurement) allowed the estimate of the top quark mass at 174GeV, give or take a dozen GeV. That measurement stood the test of time, being very close to the current world average [39].

The official top quark discovery came one year later [40, 41], when the addition of the first part of the 1994-1995 dataset allowed a crucial reduction of statistical uncertainties. As per previous agreements, CDF announced the submission of its result with two weeks of advance to DZERO. The two experiments could then independently produce observation-level effects in their data in April 1995. With that observation the picture of hadronic constituents of matter was finally completed2.

8. Conclusions

In this brief recollection I have highlighted how the path to a definitive proof of the fundamental correctness of Murray Gell-Mann’s idea of quarks as the basic building blocks of hadronic matter, and the subsequent completion of the inventory of those constituents, was riddled with difficulties and missteps but also rich with brilliant intuitions. Gell-Mann’s intuition itself could only originate from the information yielded by over a decade spent by experiments at particle accelerators in the painstaking collection of measurement of the properties of particle reactions; similarly, precise measurements of kaon branching fractions were at the roots of the ideas of a fourth quark and of a third generation of matter fields. One is thus led to observe that groundbreaking theoretical ideas are often fueled by experimental input.

For the above reason, it is in my opinion misguided to argue against continuing our quest of the unknown at the high-energy frontier with new, higher-energy particle accelerators on the grounds that ‘there are insufficient theoretical arguments’ for the existence of new physics effects at experimental reach. If we look back, we see how progress in our understanding of the fundamental properties of matter in the past century was largely driven by empirical investigation.

2 It would take a few more years before the announcement of the direct observation of tau neutrinos also completed the table of fundamental leptons [42].
ORCID iDs

Tommaso Dorigo @ https://orcid.org/0000-0002-1659-8727

References

[22] Salam A and Ward J C 1959 Weak and electromagnetic interactions N. Cim. 11 568
[35] Galtieri A private communication
[44] Galtieri A private communication