On the hadron mass decomposition

Cédric Lorcé

Centre de Physique Théorique, École polytechnique, CNRS, Université Paris-Saclay, 91128 Palaiseau, France

Received: 11 December 2017 / Accepted: 13 January 2018 / Published online: 10 February 2018
© The Author(s) 2018. This article is an open access publication

Abstract We argue that the standard decompositions of the hadron mass overlook pressure effects, and hence should be interpreted with great care. Based on the semiclassical picture, we propose a new decomposition that properly accounts for these pressure effects. Because of Lorentz covariance, we stress that the hadron mass decomposition automatically comes along with a stability constraint, which we discuss for the first time. We show also that if a hadron is seen as made of quarks and gluons, one cannot decompose its mass into more than two contributions without running into trouble with the consistency of the physical interpretation. In particular, the so-called quark mass and trace anomaly contributions appear to be purely conventional. Based on the current phenomenological values, we find that in average quarks exert a repulsive force inside nucleons, balanced exactly by the gluon attractive force.

1 Introduction

According to the standard model of particle physics, the masses of almost all known elementary particles are generated through the Brout–Englert–Higgs (BEH) mechanism. The current light quark masses obtained from such a mechanism correspond however to only about 1% of the nucleon mass. Therefore, the mass of the ordinary matter around us essentially finds its origin in the strong interactions which confine quarks and gluons inside hadrons, and not the BEH mechanism [1,2].

Understanding the origin of the hadron mass in Quantum Chromodynamics (QCD) represents a formidable challenge owing to the relativistic, quantum and non-perturbative nature of the problem. The QCD lagrangian in the chiral limit appears to be scale invariant at the classical level. This implies in particular that all the hadron masses should vanish in that limit. Scale invariance is however an anomalous symmetry, in the sense that it appears to be broken at the quantum level by radiative corrections. This gives rise, via dimensional transmutation, to a dimensionful parameter \( \Lambda_{QCD} \approx 0.2 \text{ GeV} \) in the theory [3,4], and therefore to a non-trivial spectrum.

Lattice QCD calculation of hadron masses obtained from the analysis of correlations functions in Euclidean time is in remarkable agreement with the experimental spectrum [5,6]. Unfortunately, this method gives little insight on how these masses arise from quark and gluon contributions.

In order to address the question of the origin of the hadron mass, one should rather start from the energy-momentum tensor (EMT) of QCD. At the classical level, it reads

\[
T_{\mu\nu} = \bar{\psi} \gamma^\mu \frac{i}{2} \not{D}^\nu \psi - G^{a\mu\nu} G_{a}^{\mu\nu} + \frac{1}{4} \eta^{\mu\nu} G^2,
\]

(1)

where \( \psi \) is the quark field, \( G^{a\mu\nu} \) is the gluon field strength, \( \not{D}^\mu = \not{\partial} - \not{A}^a t^a \) is the symmetric non-abelian covariant derivative, and \( \eta^{\mu\nu} \) is the Minkowski metric. Note that, as argued e.g. in [7,8], there is no need in Particle Physics to require the EMT to be symmetric in its Lorentz indices. This aspect will however not affect the purpose of the present paper and will therefore not be discussed further.

An important property of the total EMT is that it is conserved \( \partial_\mu T^{\mu\nu} = 0 \), which is a consequence of the invariance of the theory under space-time translations. At the quantum level, it does not therefore require any global renormalization [9].

The trace of the renormalized EMT is given by [10–15]

\[
T_\mu = \frac{\beta (g)}{2g^2} G^2 + (1 + \gamma_m) \bar{\psi} m \psi,
\]

(2)

where \( m \) is the quark mass matrix and \( \gamma_m \) is its anomalous dimension. The first term is known as the trace anomaly, which is a pure quantum effect as indicated by the \( \beta \) function factor. In the chiral limit \( m \to 0 \), this term prevents the trace of the EMT to vanish, and hence prevents QCD to become a scale-invariant theory with trivial spectrum.
Mass decompositions can be obtained from the expectation value of the EMT. In the next Section, we briefly present the two standard decompositions used in the literature, and point out a couple of issues regarding their physical interpretation. To the best of our knowledge, some of these issues have not been addressed before, in particular regarding pressure effects. Based on the semiclassical picture, we propose in Sect. 3 a new decomposition free of these issues. Then we discuss in Sect. 4 the new picture in more detail and we estimate the various contributions using our current phenomenological knowledge. Finally, we summarize our results in Sect. 5.

2 A critique of the standard decompositions

2.1 Trace decomposition

From Poincaré invariance, the forward matrix elements of the total EMT in a hadron state with momentum \( P \) and spin \( j \leq \frac{1}{2} \) reads [16]

\[
\langle P|T^{\mu\nu}(0)|P \rangle = 2 P^\mu P^\nu, \tag{3}
\]

where the state is normalized according to

\[
\langle P'|P \rangle = 2 P^0 (2\pi)^3 \delta^3(P' - P). \tag{4}
\]

The hadron mass \( M \) can then be defined via the trace of the EMT [16–23]

\[
2M^2 = \langle P|T^{\mu\mu}(0)|P \rangle. \tag{5}
\]

This expression is quite appealing owing to its manifest Poincaré invariance. In particular, it relies on the identity \( P^2 = M^2 \), which defines mass as the square root of one of the Casimir operators associated with the Poincaré group.

Using the explicit expression (2) for the trace of the EMT operator

\[
2M^2 = \langle P| \frac{\beta(x)}{2g} G^2 |P \rangle + \langle P|(1 + \gamma_m) \bar{\psi} m \psi |P \rangle, \tag{6}
\]

it is tempting to conclude that most of the light hadron masses comes from gluons, based on the fact that the second term on the right hand side is known to be a rather small contribution (except for the pion) [17,24–27]. Note that this partonic picture applies whenever the renormalization scale is much larger than 2 GeV. For a renormalization scale below 0.5 GeV, the picture may look quite different and closer to a non-relativistic picture where the (effective) constituents are quite massive and the (residual) binding energy is small and negative [25,26].

The physical interpretation of the origin of hadron mass based on Eq. (5) suffers however from a couple of caveats. Indeed, it seems a bit odd from a physical point of view that the definition of mass involves the trace of the EMT only at a single point. Moreover, the precise form of Eq. (3), and hence of Eq. (5), depends on how hadron states are normalized. A proper physical interpretation should in principle not depend on our conventions for the normalization of states. It is therefore more natural to define mass via the normalized expectation value of some spatially extended extended operator

\[
\langle O \rangle = \frac{\langle P| \int d^3r \ O(r)|P \rangle}{\langle P|P \rangle}. \tag{7}
\]

It is understood that such an expectation value contains implicitly the subtraction of the corresponding vacuum matrix element. Note in particular that the precise form of Eq. (3) was fixed by the requirement that \( \langle T^{0\mu} \rangle = P^\mu \). It follows that the proper expectation value of the trace of the EMT reads

\[
\langle T^{\mu\mu} \rangle = M^2 / P^0 \tag{8}
\]

independently of how states are normalized. The presence of the \( P^0 \) factor indicates that this expectation value is frame dependent.

In our opinion, frame dependence is not really a problem since physical quantities often find a simple interpretation only in some particular set of reference frames. For instance, using the light-front form of dynamics [28], Eq. (8) becomes \( \langle T^{\mu\mu} \rangle = M^2 / P^+ \) which can be interpreted in the symmetric frame (defined by \( P^+ = 0_+ \)) as twice the light-front energy \( \langle T^{\mu\mu} \rangle |P^+ = 0_+ = 2P^- \). Frame dependence\(^1\) seems to be a generic feature associated with many decompositions of a system. If one insists on manifest Lorentz invariance,\(^2\) it turns out that Lorentz non-invariant quantities defined in a given frame can be formally expressed as Lorentz scalars [7,29]. By construction, the latter will give the same result in any frame, and will coincide with the original non-invariant quantities in the given frame. Take for example the energy \( p^0 \) which is obviously not a Lorentz-invariant quantity. The rest-frame energy though can be written as a Lorentz scalar \( p^0|_{\text{rest}} = p^2 / m \). More generally, the energy in a given reference frame determined by the time-like unit four-vector \( u^\mu \) is given by \( p \cdot u \). For the rest frame, we simply have

\(^1\) Note that in the light-front form of dynamics, boost invariance is sometimes incorrectly called frame independence.
\(^2\) Sometimes people consider that physical quantities have to be Lorentz invariant, and so they conclude that decompositions are often unphysical and hence uninteresting or irrelevant. Interestingly, this seems to be the dominant thought regarding diffeomorphism invariance in General Relativity, but alternative descriptions based on the concept of vielbein are more flexible on the matter. We feel that this is too strict a requirement, for that many measurable quantities like e.g. energy and spin would be considered as unphysical.
\[ u^\mu = p^\mu / m. \] Other typical examples are the proper time and length which are Lorentz-invariant quantities interpreted respectively as the time and length measured in the rest frame of the system. In other words, it is in principle always possible to express quantities in a Lorentz-invariant form, but the price to pay is that the physical interpretation necessarily singles out a privileged frame. In the present case, the frame dependence of Eq. (8) originates from the fact that we are considering the operator \[ \int \hat{T}_\mu^\nu(r), \] which is not Lorentz invariant unlike \( T_0^0(0) \). It is however in line with the standard expectation that mass effects in hadronic matrix elements should become negligible in the ultra-relativistic regime \( P^0 \gg M \).

If one insists on preserving manifest Lorentz invariance, one can alternatively consider the integral over the proper volume \[ \int d^3r \hat{T}_\mu^\nu(r). \] Frame dependence will disappear from the formal expressions but it will remain in their physical interpretation.

Beside the question of frame dependence, a more important point we would like to stress is that the physical interpretation should be directly based on the quantum operator. Providing an interpretation at the level of the matrix element may be misleading. For example, in Eq. (5) the hadron mass

\[
\langle \bar{P} T_{\mu\nu}(0) | P \rangle = 2 \left( p^\mu p^\nu - \frac{1}{3} \eta^{\mu\nu} M^2 \right),
\]

and are renormalization scale independent, as required by Lorentz symmetry. In the Hamiltonian formalism, the mass of a system is defined as the total energy in the rest frame. It then follows that

\[
\langle \bar{M}^{00} \rangle |_{P=0} = \frac{3}{4} M, \quad \langle \hat{M}^{00} \rangle |_{P=0} = \frac{1}{4} M.
\]

In other words, the trace of the EMT appears to contribute only to 25% of the hadron mass.

Ji proceeded with a decomposition of the traceless part into quark and gluon contributions, and of the trace part into quark mass and trace anomaly contributions

\[
\hat{T}_{\mu\nu}^{\text{trace}} = \hat{T}_{\text{q}}^{\mu\nu} + \hat{T}_{\text{g}}^{\mu\nu} + \hat{T}_{\text{m}}^{\mu\nu} + \hat{\eta}^{\mu\nu}. \tag{13}
\]

The corresponding individual forward matrix elements are not protected by Lorentz symmetry

\[
\langle \bar{P} \hat{T}_{\text{q}}^{\mu\nu}(0) | P \rangle = 2 a(\mu^2) \left( p^\mu p^\nu - \frac{1}{3} \eta^{\mu\nu} M^2 \right),
\]

and so the coefficients \( a(\mu^2) \) and \( b(\mu^2) \) depend generally on the renormalization scale \( \mu \). It follows from Eq. (13) that the Hamiltonian operator can naturally be decomposed as

\[
H = H_q' + H_g' + H_m' + H_a', \tag{18}
\]

where

\[
H_q' = \int d^3r \bar{T}_q^{00}(r), \tag{19}
\]

\[
H_g' = \int d^3r \bar{T}_g^{00}(r). \tag{20}
\]

3 The general trace part is defined as \( \hat{\eta}^{\mu\nu} = \hat{T}_{\mu\nu} + c \hat{\eta}_{\mu\nu} \) with \( c \) an arbitrary constant.

2.2 Ji’s decomposition

About 20 years ago, Ji proposed a decomposition of the hadron mass analogous to the virial theorem for a harmonic oscillator and the hydrogen atom [31,32], see also [33]. As a first step, he decomposed the QCD EMT into traceless and trace parts

\[
\hat{T}_{\mu\nu} = \hat{T}_{\mu\nu}^{\text{traceless}} + \hat{T}_{\mu\nu}^{\text{trace}}
\]

with \( \hat{T}_{\mu\nu}^{\text{trace}} = \frac{1}{2} \eta^{\mu\nu} T_0^0 \). This choice is of course not unique, but it is well suited for treating the trace anomaly contribution.
put in a covariant form by considering the Lorentz-invariant quantity \((T^0_0 u_\mu)\), where \(u^\mu \equiv P^\mu / M\) is interpreted as the hadron four-velocity. Although this form is frame independent, its physical interpretation becomes simple only in the rest frame, where \(u^\mu = (1, 0)\). For a massless state, since there is no rest frame, one can consider instead the energy decomposition in any frame. In this case, there is no contribution from the trace part since \(P^2 = 0\) and so only the coefficient \(\alpha (\mu^2)\) is needed.\(^5\)

The actual problem with Ji’s decomposition is that the separation of the EMT into traceless and trace parts is not inconsequential for the physical interpretation of the individual contributions.\(^6\) Namely, although \(T^{00}\), \(\tilde{T}^{00}\) and \(\tilde{T}^{00}\) all have the dimension of energy densities, they actually correspond to different thermodynamic potentials as we will show in the following. By focusing on the \(\mu = v = 0\) component in the rest frame, Ji’s decomposition does not make any distinction between the Lorentz tensors \(P^\mu P^\nu\) and \(M^2 \eta^{\mu\nu}\), and therefore disregards pressure effects.

To sum up, although all terms in Ji’s decomposition (27) can formally be defined and evaluated, they cannot however be interpreted as pure mass contributions. In the next section, we will explain how to achieve a proper mass decomposition based on a more covariant treatment using the semiclassical picture.

3 A new decomposition

Consider some decomposition of the EMT at the operator level \(T^{\mu\nu} \equiv \sum_j T^{\mu\nu}_j\). From the general parametrization of the matrix element of a rank-two tensor in a state with mass \(M\) and spin \(j \leq \frac{1}{2}\) [7, 39, 40], we obtain in the forward limit \(\langle P | T^{\mu\nu}_j (0) | P \rangle = 2 P^\mu P^\nu A_j (0) + 2 M^2 \eta^{\mu\nu} \tilde{C}_j (0)\), (33)\n
where \(A_j (0)\) and \(\tilde{C}_j (0)\) are two energy-momentum form factors evaluated at zero momentum transfer and depending on the renormalization scale \(\mu\). They are related to Ji’s coefficients as follows

\[
\alpha_i \equiv A_j (0), \quad \beta_i \equiv A_j (0) + 4 \tilde{C}_j (0). \quad (34)
\]

\(^5\) Note that one could still formally use Ji’s decomposition based on the covariant quantities \(\lim_{\vec{P} \to 0} \langle T^{\mu_0}_\mu \vec{P}_\mu / \vec{P}^2 \rangle\) or \(\partial \langle T^{\mu_0}_\mu \vec{P}_\mu \rangle / \partial \vec{P}^2|_{\vec{P}^2=0}\), which are in the spirit of Ji’s remark [32] that although the overall scale is essentially determined by the anomaly in the light hadron sector, the relative magnitudes of the various contributions reflect essential aspects of the underlying quark-gluon dynamics in the non-perturbative regime.

\(^6\) Paying attention not to introduce spurious contributions that sum up to zero in a decomposition is a general problem. For example, the proper definition of angular momentum at the level of spatial distribution has recently been discussed in detail in [38], where the problem was solved by treating with care all contributions that vanish under integration.
Four-momentum conservation implies that
\[ \sum_i A_i(0) = 1, \quad \sum_i \dot{C}_i(0) = 0, \]  
(35)
as one can see from Eq. (3).

In order to identify the physical meaning of \( T_i^{\mu \nu} \), we consider its expectation value averaged over the hadron volume \( V = M V / P_0 \), with \( V \) denoting the hadron proper volume,
\[ \frac{1}{V} \langle T_i^{\mu \nu} \rangle = \frac{\langle P | T_i^{\mu \nu}(0) | P \rangle}{2 M V}. \]  
(36)

In the hadron rest frame, we find
\[ \frac{1}{V} \langle T_i^{\mu \nu} \rangle |_{p=0} = \begin{pmatrix} \epsilon_i & p_i \\ p_i & p_i \end{pmatrix}, \]  
(37)
where
\[ \epsilon_i = [A_i(0) + \dot{C}_i(0)] \frac{M}{V} \quad \text{and} \quad p_i = -\dot{C}_i(0) \frac{M}{V}. \]  
(38)

are naturally interpreted as partial proper (internal) energy density and isotropic pressure, respectively. Note that \( \epsilon_i \) and \( p_i \) are average quantities over the hadron proper volume, and hence are proportional to the average proper energy density of the hadron \( M / V \). In a generic frame, we can write
\[ \frac{1}{V} \langle T_i^{\mu \nu} \rangle = (\epsilon_i + p_i) u^\mu u^\nu - p_i \eta^{\mu \nu} \]  
(39)
which has the same structure as partial internal energy density and isotropic pressure, respectively. Each continuum is then described by its own (partial) energy density \( \epsilon_i \), and pressure \( p_i \).

Remembering that \( \epsilon_i \) and \( p_i \) are average quantities over the proper volume, the thermodynamic potentials are then simply defined as
\[ U_i = \epsilon_i V = [A_i(0) + \dot{C}_i(0)] M, \quad H_i = h_i V = A_i(0) M, \]  
(40)
where \( U_i \) is the partial internal energy, \( H_i = U_i + W_i \) is the partial enthalpy, and \( W_i = p_i V \) is the partial pressure–volume work. The latter quantity represents the energy associated with the occupation of the volume \( V \) by the subsystem \( i \) pushing \( (p_i > 0) \) or pulling \( (p_i < 0) \) the rest of the system.

Since the internal energy represents the sum of all the kinetic and potential energies, the proper mass decomposition we were looking for is
\[ M = \sum_i U_i. \]  
(42)

Moreover, the total pressure–volume work should vanish for a stable system
\[ \sum_i W_i = 0. \]  
(43)
These two conditions are naturally equivalent to Eq. (35).

4 Discussion

In the previous section, we observed that the generic form of the EMT for a hadron with spin \( j \leq \frac{1}{2} \) is characterized by two Lorentz scalar quantities similarly to that of an element of perfect fluid. We are of course not claiming that hadrons consist of a set of perfect fluids,\(^7\) let alone that a hydrodynamical description is quite difficult to justify in this context. Forward matrix elements of the EMT, like the ones we considered, allow us to determine only static mechanical properties of the state. In the semiclassical limit, these matrix elements can effectively be thought of as a continuum description averaged over time \([34,44,45]\). We simply adopted the terminology of Continuum Mechanics to identify the physical meaning of the various energy-momentum form factors.

Decomposing the EMT into various contributions \( T^{\mu \nu} = \sum_i T_i^{\mu \nu} \) amounts in practice to an effective (time-averaged) coupled multifluid picture of the hadron \([46]\), where each species of the mixture is regarded as a separate continuum coexisting with the continuums made up of other species. Each continuum is then described by its own (partial) energy density \( \epsilon_i \) and pressure \( p_i \). Since these quantities are averaged over the hadron proper volume, they trivially obey a barotropic equation of state for species \( i \)
\[ p_i = w_i \epsilon_i. \]  
(44)

Depending on the value of the parameter \( w_i \), this equation of state describes an element of effective perfect fluid from phantom to ekpyrotic matter \([47,48]\).

\(^7\) Similarly, scalar fields in General Relativity are often mimicked by effective perfect fluids, see e.g. \([42,43]\).
Hadrons are bound states of two types of constituents, namely quarks and gluons. We can therefore decompose the QCD EMT into quark and gluon contributions

\[ T^{\mu\nu} = T_q^{\mu\nu} + T_g^{\mu\nu}, \]

where

\[ T_q^{\mu\nu} = \psi^* \gamma^\mu \frac{1}{2} \overleftarrow{D^\nu} \psi + \frac{1}{4} \eta^{\mu\nu\rho\lambda} \overleftrightarrow{W}_\rho \psi, \]

\[ T_g^{\mu\nu} = -G^{\mu\nu\lambda} G^{\alpha\lambda} \lambda + \frac{1}{2} \eta^{\mu\nu} \left( 1 + \frac{b(\psi)}{28} \right) G^2, \]

although there is some arbitrariness in doing so owing to quark-gluon interactions [7]. Here, the quark mass term is part of \( T_q^{\mu\nu} \) and the trace anomaly is part of \( T_g^{\mu\nu} \). Accordingly, it seems therefore natural to adopt an effective coupled two-fluid picture where \( T_q^{\mu\nu} \) and \( T_g^{\mu\nu} \) are treated as separate EMT and hence described in terms of their own energy densities \( \varepsilon_{q,g} \) and pressures \( p_{q,g} \).

Let us now come back to Ji’s decomposition. Looking at the expressions (29)–(32), we observe that the individual contributions to the hadron mass correspond to different linear combinations of the coefficients \( a \) and \( b \). In the effective coupled two-fluid picture, they correspond therefore to different linear combinations of energy density and pressure owing to Eq. (40). This is of course not physically acceptable, in the sense that one is adding apples and oranges. It is however mathematically correct since the sum over species gives at the end just the hadron mass \( M \), once again thanks to the stability constraint \( \sum W_i = 0 \). The only way to make physical sense out of Ji’s decomposition within the effective coupled two-fluid picture is to define the quark mass and trace anomaly contributions \( \varepsilon_{m,a} \) from the onset, and to split the quark and gluon energies as follows

\[ \varepsilon_q = \varepsilon_{q-m} + \varepsilon_m, \quad \varepsilon_g = \varepsilon_{g-a} + \varepsilon_a, \]

with \( \varepsilon_{q-m} \equiv \varepsilon_q - \varepsilon_m \) and \( \varepsilon_{g-a} \equiv \varepsilon_g - \varepsilon_a \), while keeping the pressure part unchanged. This is consistent, but the price to pay is a loss of covariance since different components of the EMT are not treated in the same way.

What Ji did actually in Eq. (13) is a covariant decomposition of the QCD EMT into four parts. He considered with Eqs. (19)–(22) that each quantity \( T_q^{00}, T_g^{00}, T_m^{00}, \) and \( T_a^{00} \) should represent an energy density, and hence implicitly adopted an effective coupled four-fluid picture of the hadron. The problem with such a description is that e.g. gluons carrying kinetic and potential energies are considered as different from those involved in the trace anomaly, in the sense that they are effectively treated as two distinct continuums with their own equations of state. Our opinion is that this is not acceptable from a physical point of view. Since hadrons are

---

8 In principle, one can construct infinitely many decompositions of the form \( M = \sum_i (U_i + c W_i) \) with \( c \) an arbitrary constant.

9 Note that in the context of teleparallel gravity, the concepts of gravitational energy and pressure can be defined, see e.g. [49–51].
composed of quarks and gluons, an effective description of hadrons in terms of only two continuums is more natural.

4.2 Virial decomposition

In the effective coupled two-fluid picture, the hadron mass is decomposed into quark and gluon internal energies \( M = U_q + U_g \), where \( U_{q,g} = \langle T_{q,g}^{00} \rangle |_{p=0} \) with \( T_{q,g}^{00} \) given by Eqs. (49) and (50). The question now is whether one can further decompose into parton kinetic and potential energies, quark mass and trace anomaly contributions, similarly to Ji’s decomposition.

Let us start with the gluon sector. Since the only contribution to the trace of the gluon EMT comes from the trace anomaly, it seems natural to write

\[
T_{g}^{\mu \nu} = \tilde{T}_{g}^{\mu \nu} + \hat{T}_{g}^{\mu \nu}.
\]

For the quark sector, we resuffle the quark mass terms after the separation into traceless and trace parts

\[
T_{q}^{\mu \nu} = \tilde{T}_{q}^{\mu \nu} + \hat{T}_{q}^{\mu \nu},
\]

where

\[
\tilde{T}_{q}^{\mu \nu} = \tilde{T}_{q}^{\mu \nu} - c \hat{T}_{q}^{\mu \nu}, \quad \hat{T}_{q}^{\mu \nu} = (1 + c) \hat{T}_{q}^{\mu \nu}.
\]

The constant \( c \) is then chosen such that \( \tilde{T}_{q}^{00} \) takes the form \( \psi^{\dagger}(-i \mathbf{D} \cdot \alpha)\psi \) upon using the QCD equations of motion, leading us to

\[
c = \frac{3}{1 + \gamma_m}.
\]

Now, contrary to Ji, we will not treat \( \tilde{T}_{q}^{\mu \nu} \) and \( \hat{T}_{q}^{\mu \nu} \) as separate EMT, but rather as mere parts of the EMT \( T_{q}^{\mu \nu} \). We then have

\[
\langle \tilde{T}_{q}^{00} \rangle |_{p=0} = \frac{3}{4} (U_i + W_i), \quad \langle \tilde{T}_{q}^{00} \rangle |_{p=0} = \frac{1}{4} (U_i - 3W_i).
\]

In other words, we see that traceless parts contribute to \( \frac{3}{4} \) and trace parts to \( \frac{1}{4} \) of the internal energy content. Summing over all the species \( i \) naturally leads to Eq. (12). Keeping only the internal energy contributions leads to the following finer decomposition of the hadron mass in the effective coupled two-fluid picture

\[
M = \tilde{U}_q + \hat{U}_q + \tilde{U}_g + \hat{U}_g
\]

with

\[
\tilde{U}_q = \frac{3}{4} \frac{\gamma_m}{1 + \gamma_m} U_q, \quad \hat{U}_q = \frac{1}{4} \frac{4 + \gamma_m}{1 + \gamma_m} U_q,
\]

\[
\tilde{U}_g = \frac{3}{4} U_g, \quad \hat{U}_g = \frac{1}{4} U_g.
\]

For comparison, Ji’s decomposition reads explicitly

\[
M_q = \langle \tilde{T}_{q}^{00} \rangle |_{p=0} = \tilde{U}_q + \frac{3}{4} \frac{4 + \gamma_m}{1 + \gamma_m} W_q,
\]

\[
M_m = \langle \hat{T}_{q}^{00} \rangle |_{p=0} = \hat{U}_q - \frac{3}{4} \frac{4 + \gamma_m}{1 + \gamma_m} W_q,
\]

\[
M_g = \langle \hat{T}_{q}^{00} \rangle |_{p=0} = \tilde{U}_g + \frac{3}{4} W_g,
\]

\[
M_m = \langle \tilde{T}_{q}^{00} \rangle |_{p=0} = \tilde{U}_g - \frac{1}{4} W_g,
\]

and differs from our finer decomposition by a reshuffling of the energy contributions within the quark and gluon sectors, separately.

We refrain from interpreting \( \tilde{U}_q \) as quark mass contribution and \( \hat{U}_g \) as trace anomaly contribution. Indeed, in the classical limit \( \gamma_m \to 0 \) we have \( \tilde{U}_q = U_q \) which should also include quark kinetic and potential energies. Moreover, the contribution \( \hat{U}_g \) does not vanish when the trace anomaly is set to zero \( \langle \hat{T}_{g}^{00} \rangle = 0 \). By keeping only the internal energy contributions, we lost the direct connection with the matrix elements and hence a simple physical interpretation. At best, our finer hadron mass decomposition (57) can be seen as some sort of virial decomposition [31–33].

A similar treatment can be applied to the pressure–volume work. In the effective coupled two-fluid picture, the partial pressure–volume works of quarks and gluons satisfy \( 0 = W_q = W_g \), where \( W_{q,g} = \langle T_{q,g}^{33} \rangle |_{p=0} \). Treating once again \( \tilde{T}_{i}^{33} \) and \( \hat{T}_{i}^{33} \) as separate EMT, we get

\[
\langle \tilde{T}_{i}^{33} \rangle |_{p=0} = \frac{1}{4} (W_i + U_i), \quad \langle \hat{T}_{i}^{33} \rangle |_{p=0} = \frac{1}{4} (3W_i - U_i).
\]

In this case, traceless parts contribute to \( \frac{1}{4} \) and trace parts to \( \frac{3}{4} \) of the pressure–volume work. Summing over all the species \( i \), we obtain the analogue of Eq. (12) for pressure–volume work

\[
\langle \tilde{T}_{33} \rangle |_{p=0} = \frac{1}{4} M, \quad \langle \hat{T}_{33} \rangle |_{p=0} = -\frac{1}{4} M.
\]

Keeping only the pressure–volume work contributions leads to the following finer decomposition of the hadron stability constraint in the effective coupled two-fluid picture

\[
0 = \tilde{W}_q + \tilde{W}_q + \tilde{W}_g + \hat{W}_g
\]

with

\[
\tilde{W}_q = -\frac{1}{4} \frac{8 - \gamma_m}{1 + \gamma_m} W_q, \quad \tilde{W}_q = \frac{3}{4} \frac{4 + \gamma_m}{1 + \gamma_m} W_q,
\]

\[
\tilde{W}_g = \frac{1}{4} W_g, \quad \tilde{W}_g = \frac{3}{4} W_g.
\]
For comparison, a similar decomposition in the effective coupled four-fluid picture reads

\[ 0 = W_q + W_m + W_g + W_a \]  

(67)

with

\[ W_q = \langle \tilde{T}_{q}^{33} \rangle |_{p=0} = \tilde{W}_q + \frac{4 + 2n}{3 + 2n} U_q, \]  

(68)

\[ W_m = \langle \tilde{T}_{g}^{33} \rangle |_{p=0} = \tilde{W}_g - \frac{4 + 2n}{3 + 2n} U_q, \]  

(69)

\[ W_g = \langle \tilde{T}_{g}^{33} \rangle |_{p=0} = \tilde{W}_g + \frac{4}{3} U_g, \]  

(70)

\[ W_a = \langle \tilde{T}_{g}^{33} \rangle |_{p=0} = \tilde{W}_g - \frac{4}{3} U_g, \]  

(71)

and differs from our finer decomposition by a reshuffling of the work contributions within the quark and gluon sectors, separately. Once again, we refrain from interpreting \( \tilde{W}_q \) as quark mass contribution and \( \tilde{W}_g \) as trace anomaly contribution. At best, Eq. (65) can be seen as some sort of virial decomposition.

The above finer decompositions are not based on the nature of the constituents but on Lorentz symmetry and this has a dramatic impact on the physical picture. Indeed, we were not able find a simple physical interpretation for the individual contributions to the decompositions (57) and (65), because the effective coupled two-fluid picture forced us to discard some of the terms in the matrix elements for consistency. In order to avoid this, one has to adopt a picture where the number of effective fluids is at least equal to the number of contributions. In the present case, we need at least an effective four-fluid picture, leading us back directly to Ji’s decomposition. We have however already argued that the four-fluid picture is not natural from the physical point of view, since gluons carrying kinetic and potential energies are the same as those involved in the trace anomaly. Similarly, quarks carrying kinetic and potential energies are the same as those characterized by mass \( m \).

Another indication that a decomposition based on Lorentz symmetry is not natural from the physical point of view is the following. Any EMT with the structure (37) and generic equation of state \( p_i = w_i e_i \) can be represented as a particular combination of two EMT with fixed equations of state. This can be seen as some sort of decomposition onto a basis. For example, through a decomposition of the EMT into traceless and trace parts \( T_{\mu \nu i} = \tilde{T}_{\mu \nu i} + \tilde{T}_{\mu \nu i} \), the set \( \{e_i, p_i\} \) can formally be replaced by \( \{\tilde{e}_i, \tilde{p}_i\} \) with \( \tilde{e}_i = \frac{3}{4} (e_i + p_i) \) and \( \tilde{p}_i = \frac{1}{4} (e_i - 3p_i) \), since traceless EMT \( \tilde{T}_{\mu \nu i} \) are characterized by the equation of state \( \tilde{p}_i = \frac{1}{4} \tilde{e}_i \) and pure trace EMT \( \tilde{T}_{\mu \nu i} \) are characterized by \( \tilde{p}_i = -\tilde{e}_i \). In Ji’s decomposition, the gluon contribution is indeed divided into kinetic and potential energies treated as a pure radiation \( w_g = \frac{1}{4} \), and trace anomaly treated as a cosmological constant \( w_a = -1 \). Since the choice of a basis is not unique, the decomposition based on Lorentz symmetry is purely conventional, and hence artificial. In fact, Lorentz symmetry has already been used to provide a physical interpretation of the various components of the EMT, namely by distinguishing in our case energy density \( e_i \) from pressure \( p_i \) in Eq. (37). One can then easily understand why using Lorentz symmetry again to perform a decomposition amounts to choosing an arbitrary basis of EMT with fixed equations of state not determined by the physics of the problem.

To summarize, although a decomposition of the hadron mass or pressure–volume work into four contributions can formally be achieved in the effective coupled two-fluid picture, the individual terms cannot be interpreted in a simple way. If the hadron is seen as made of quarks and gluons, the only unambiguous decompositions are

\[ M = U_q + U_g, \quad 0 = W_q + W_g. \]  

(72)

Any further decomposition based on Lorentz symmetry, like e.g. a separation into traceless and trace parts (or variations thereof like \( U_q = \tilde{U}_q + \bar{U}_q = M_q + M_m \) and \( U_g = \tilde{U}_g + \bar{U}_g = M_g + M_a \)), is artificial and purely arbitrary. The only physically acceptable further decompositions are those based on constituent properties, like e.g. flavor or polarization [52, 53].

### 4.3 Phenomenology

In practice one cannot extract directly from an experiment the matrix elements of the EMT owing to the weakness of the gravitational force. However, all that is needed to characterize the quark and gluon EMT (33) are the form factors \( A_{q,g}(0) \) and \( \tilde{C}_{q,g}(0) \) in the forward limit, and these can be extracted from other more directly accessible physical amplitudes thanks to operator identities. Using the energy-momentum sum rules (35), we can reduce this set to e.g. the quark form factors \( A_q(0) \) and \( \tilde{C}_q(0) \), or equivalently Ji’s coefficients \( a \) and \( b \) owing to Eq. (34). In the following, we will consider only the proton case and fix the renormalization scale to \( \mu = 2 \, \text{GeV} \).

The parameter \( a \), which is also interpreted as the average fraction of hadron momentum carried by quarks, can be extracted from deep-inelastic lepton-proton scatterings. A recent global analysis with leading-order parametrization obtained \( a = 0.546 \pm 0.005 \) [54]. The parameter \( b \) is related to the scalar charge of the proton and has been estimated to \( b = 0.113 \pm 0.010 \) by Gao et al. [2], based on a recent determination of the pion–nucleon \( \sigma \)-term [55], a recent lattice calculation of the strangeness content [56] and neglecting the heavy quark contributions. It has also been suggested to extract the parameter \( b \) from the trace anomaly using quarkonium-hadron scattering close to threshold [21,57]. For completeness, the anomalous quark mass dimension is
approximately given by $\gamma_m \approx -0.15$ for $n_f = 3$ active flavors [58]. The various decompositions obtained with these values are depicted in Figs. 1, 2, 3 and 4.

In Fig. 1 we represent the trace decomposition given by Eq. (6) and determined only by the parameter $b$. As already discussed, it is largely dominated by the gluon contribution and is at the origin of the claim that most of the nucleon mass comes from gluons [17, 24–27]. The trace of the EMT being given by

$$T_{\mu\mu} = T_{00} - \sum_j T_{jj},$$

what the quark and gluon contributions in Eq. (6) do actually represent are the combinations $U_q, g - 3W_{q, g}$. Since the total pressure–volume work vanishes, one artificially suppresses one of the contributions in favor of the other with pressure effects. As we will see below, it turns out that $W_q = -W_g > 0$ which explains why the quark contribution appears to be much smaller than the gluon contribution. Because of these pressure effects, the sole trace of the QCD EMT does not provide enough information to determine the actual quark and gluon contributions to the hadron mass.

Ji’s decomposition (27) is represented in Fig. 2. The $M_q$ and $M_m$ contributions slightly differ from the ones in the pie chart drawn in [2] because we did not neglect the anomalous quark mass dimension $\gamma_m$. All the information about the hadron EMT in the forward limit, parametrized by Ji’s coefficients $a$ and $b$, is included in this decomposition. As argued in the previous sections, a decomposition of the hadron mass into four contributions is however artificial. In the effective coupled two-fluid picture, the four contributions appear to be combinations of internal energies and pressure–volume works

$$M_q = \frac{3}{4} \frac{4 + \gamma_m}{4 + \gamma_m} \left( \frac{\gamma_m}{4 + \gamma_m} U_q + W_q \right),$$

$$M_m = \frac{1}{4} \frac{4 + \gamma_m}{4 + \gamma_m} \left( U_q - 3W_q \right),$$

$$M_g = \frac{3}{4} \left( U_g + W_g \right),$$

$$M_a = \frac{1}{4} \left( U_g - 3W_g \right).$$

In order to avoid adding apples and oranges, Ji required that each term must represent some internal energy, which
amounts to adopting an effective coupled four-fluid picture. In such a picture, no new information arises from the corresponding decomposition of the pressure–volume work (67) represented in Fig. 3, since the individual contributions are directly determined by the energy ones

\[ W_q = \frac{1}{3} \left( M_q + \frac{12}{1+{\gamma_{m}}} M_m \right), \quad W_m = -M_m, \]
\[ W_g = \frac{1}{3} M_g, \quad W_a = -M_a, \quad (77) \]
as a consequence of the conventional splitting into traceless and trace parts.\(^{10}\) In the effective coupled two-fluid picture, they are explicitly given by

\[ W_q = \frac{1}{4} \frac{4+\gamma_{m}}{1+\gamma_{m}} \left( U_q - \frac{8-\gamma_{m}}{4+\gamma_{m}} W_q \right), \quad (78) \]
\[ W_m = -\frac{1}{4} \frac{4+\gamma_{m}}{1+\gamma_{m}} \left( U_q - 3 W_q \right), \quad (79) \]
\[ W_g = \frac{1}{4} \left( U_g + W_g \right), \quad (80) \]
\[ W_a = -\frac{1}{4} \left( U_g - 3 W_g \right). \quad (81) \]
The information about the energy content and the stability of the system are a priori independent. By treating the traceless and trace parts as separate EMT, Ji effectively combined both information into a single decomposition. While mathematically correct, such a decomposition is however not satisfactory from the physical point of view because gluons involved in the trace anomaly are treated as if they were distinguishable from those carrying kinetic and potential energies (they are indeed associated with different equations of state in Ji’s approach).

Since hadrons are made of quarks and gluons, it seems more natural to decompose into two contributions only. Ji’s coefficients \( a \) and \( b \) appear to characterize two different hadron properties of the hadron, namely mass and pressure–volume work. Accordingly, instead of a single decomposition of the hadron mass into four contributions like Ji’s, we propose to consider separately the decompositions of hadron mass and pressure–volume work into two contributions, given by Eq. (72) and represented in Fig. 4. We do not require any splitting into traceless and trace parts, and hence we do not fix a priori the relation between energy density and pressure. Note that all the standard energy conditions are satisfied [59,60]

null energy condition \( \epsilon_i + p_i \geq 0 \),
weak energy condition \( \epsilon_i \geq 0 \) and \( \epsilon_i + p_i \geq 0 \),
dominant energy condition \( \epsilon_i \geq |p_i| \),
strong energy condition \( \epsilon_i + p_i \geq 0 \) and \( \epsilon_i + 3p_i \geq 0 \).

\[ (82) \]

Note that the appearance of \( M_m \) in the expression for \( W_q \) comes from the reshuffling the quark mass terms between the quark traceless and trace parts (54).

5 Conclusions

We used forward matrix elements of the energy-momentum tensor to characterize static mechanical properties of hadrons. The components of such an energy-momentum tensor can be interpreted semi-classically in terms of parton energy density and pressure averaged over time and the hadron proper volume.

We showed that, because of pressure effects, the physical interpretation of the standard decompositions of the hadron mass have to be considered with care. In the trace decomposition, a pressure–volume work contribution artificially emphasizes the role played by gluons. In Ji’s decomposition, the splitting of the quark and gluon energy-momentum tensors into traceless and trace parts mixes the information about the hadron mass budget with the pressure–volume work budget. This can be understood by the fact that Lorentz symmetry is already used to provide a physical interpretation of the various components of the energy-momentum tensor. It cannot be used a second time to define separate mass contributions without modifying the physical picture. In particular, from the point of view of physical interpretation, using Ji’s decomposition amounts to treating gluons involved in the trace anomaly and those carrying kinetic and potential energies as separate entities with different equations of state, which is not physically acceptable.

Since hadrons are made of quarks and gluons, it is natural to decompose their mass into two contributions only. Any further decomposition that is not based on the properties of the constituents will be somewhat arbitrary and will mix internal energy with pressure–volume work. This mixing is mathematically harmless because the total pressure–volume work vanishes for a stable system, but it is a problem for the physical interpretation of the individual contributions.

We proposed a new picture, where the hadron mass and pressure–volume work budgets are kept separate and expressed in terms of the sole quark and gluon contributions which are physically unambiguous. In particular, unlike Ji’s decomposition we do not fix a priori the equations of state for quarks and gluons.

Finally, we quantitatively compared the different decompositions based on recent phenomenological estimates. It turned out that, as expected for highly relativistic systems, pressure–volume work contributions are of the same order of magnitude as internal energy contributions. In particular, quarks are responsible in average for the repulsive force and gluons for the attractive force inside nucleons.
Having clearly identified the pressure contributions opens many interesting applications related to compact stars. For example, determining the quark and gluon equations of state inside a nucleon may give important clues about the internal structure of compact stars.

Acknowledgements This work is a result of discussions held at the workshop “The Proton Mass: At the Heart of Most Visible Matter” at the ECT* Trento, on 3–7 April 2017. This work has been supported by the Agence Nationale de la Recherche Under the project ANR-16-CE31-0019.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP3.

References

1. M. Schumacher, Ann. Phys. 526(5–6), 215 (2014)
2. H. Gao, T. Liu, C. Peng, Z. Ye, Z. Zhao, The Universe 3(2), 18 (2015)
3. C.G. Callan Jr., S.R. Coleman, R. Jackiw, Ann. Phys. 59, 42 (1970)
4. S.R. Coleman, E.J. Weinberg, Phys. Rev. D 7, 1888 (1973)
5. S. Durr et al., Science 322, 1224 (2008)
6. R.G. Edwards et al. [Hadron Spectrum Collaboration], Phys. Rev. D 87(5), 054506 (2013)
7. E. Leader, C. Lorcé, Phys. Rept. 541(3), 163 (2014)
8. C. Lorcé, JHEP 1508, 045 (2015)
9. L.S. Brown, Ann. Phys. 126, 135 (1980)
10. R.J. Crewther, Phys. Rev. Lett. 28, 1421 (1972)
11. M.S. Chanowitz, J.R. Ellis, Phys. Lett. 40B, 397 (1972)
12. N.K. Nielsen, Nucl. Phys. B 97, 527 (1975)
13. S.L. Adler, J.C. Collins, A. Duncan, Phys. Rev. D 15, 1712 (1977)
14. J.C. Collins, A. Duncan, S.D. Joglekar, Phys. Rev. D 46, 438 (1997)
15. N.K. Nielsen, Nucl. Phys. B 210, 212 (1977)
16. R.L. Jaffe, A. Manohar, Nucl. Phys. B 337, 509 (1990)
17. M.A. Shifman, A.I. Vainshtein, V.I. Zakharov, Phys. Lett. B 78, 443 (1978)
18. J.F. Donoghue, UMHEP-284
19. M.E. Luke, A.V. Manohar, M.J. Savage, Phys. Lett. B 288, 355 (1992)
20. J.F. Donoghue, E. Golowich, B.R. Holstein, Camb. Monogr. Part Phys. Nucl. Phys. Cosmol. 2, 1 (1992) [Camb. Monogr. Part Phys. Nucl. Phys. Cosmol. 35 (2014)]
21. D. Kharzeev, Proc. Int. Sch. Phys. Fermi 130, 105 (1996)
22. M.A. Shifman, World Sci. Lect. Notes Phys. 62, 1 (1999)