Foundations of Physics (2024) 54:57
https://doi.org/10.1007/s10701-024-00795-1

REVIEW

sQFT: An Autonomous Explanation of the Interactions of Quantum Particles

K.-H. Rehren[1] · L. T. Cardoso[2] · C. Gass[3] · J. M. Gracia-Bondía[4] · B. Schroer[5] · J. C. Várilly[6]

Received: 29 May 2024 / Accepted: 2 August 2024 / Published online: 18 August 2024

Abstract
Successful applications of a conceptually novel setup of Quantum Field Theory, that accounts for all subtheories of the Standard Model (QED, Electroweak Interaction and Higgs, Yang–Mills and QCD) and beyond (Helicity 2), call for a perspective view in a broader conceptual context. The setting is "autonomous" in the sense of being intrinsically quantum. Its principles are: Hilbert space, Poincaré symmetry and causality. Its free quantum fields are obtained from Wigner's unitary representations of the Poincaré group, with only physical and observable degrees of freedom. A "quantization" of an "underlying" classical theory is not needed. It allows renormalizable perturbation theory with interactions whose detailed structure, and in some cases even the particle content, is predicted by internal consistency. The results confirm and extend observable predictions for the interactions of the Standard Model without assuming a "principle" of gauge invariance.

Keywords Quantum field theory · String-localized quantum fields · Perturbation theory · Standard Model interactions

1 The Autonomous Quantum Approach to QFT

We review "string-localized quantum field theory (sQFT)" by contrasting its conceptual ambitions with standard approaches to perturbative QFT. In later sections, we shall explain in a largely untechnical way how sQFT proceeds, and report the results it has produced by now in the context of the Standard Model of particles (SM).

Conceptual issues of QFT. Standard textbook approaches to QFT start from classical field theory plus "quantization". At the level of free fields, the strategy is to use a classical free Lagrangian and "derive" canonical commutation relations that have to be fulfilled in a covariant way. This initial step is beset with problems:

Dedicated to Detlev Buchholz on the occasion of his 80th birthday.

Extended author information available on the last page of the article

In the case of the Dirac field, the canonically conjugate momenta are functions of the field itself; and one has to replace equal-time CCR by CAR ad hoc. In the case of the Maxwell field, one has to take the unobservable gauge potential as “canonical” variables, one has to change the dynamics by a “gauge fixing term” in order to get a canonically conjugate momentum for all components of the potential, and one gets CCR that can only be represented on a Fock space with indefinite inner product (a Krein space), which contains unphysical “photon states”. We have learned how to circumvent such trouble by ingenious tricks (Dirac’s theory of constraints, the Gupta–Bleuler method to eliminate unphysical states with negative or zero norm square, and the normal-ordering prescription to make the vacuum energy finite). But we would rather prefer to avoid these first steps altogether.

When it comes to adding an interaction Lagrangian, one usually invokes gauge invariance, a postulate on unobservable gauge-dependent quantities.[1] In order to make the previous quantization methods survive the interaction, in many cases of interest one has to introduce ghost fields and promote the Gupta–Bleuler condition to BRST invariance.

Apart from its need of unphysical degrees of freedom in gauge theories, one may object against “quantization” at a more fundamental level. Classical physics is a limiting regime in a quantum world, applicable in situations where the uncertainty bound is largely exceeded. A prescription to “read back” the true theory from its limit precisely in the regime where the latter is *not* applicable, seems a hazardous endeavour. It goes a long way towards explaining the “need” for artful but artificial devices like the ones just exposed.

The autonomous quantum approach, advocated in this paper, instead starts from quantum principles: Hilbert space, unitary Poincaré symmetry and Causality. In the words of Jordan [20]: it shows how QFT can be formulated “without the crutches of classical correspondences”. Along the way, it will show that the Wightman axioms (postulating causality or anti-causality for all Wightman fields, as an extrapolation from free fields) fall too short for interacting charged fields. sQFT tells us how interacting charged quantum fields behave instead.

The autonomous approach proceeds as follows. One or several of Wigner’s unitary representations of the Poincaré group [39] for one’s favorite physical particles give rise to a Fock space, which is a Hilbert space. The generators of the Poincaré symmetry of the one-particle space lift to the Fock space, and their lifts can be expressed as integrals over densities that are already normal-ordered from the outset. Covariant free fields may be defined with Weinberg’s method [38], intertwining momentum space Fock space operators with position space local fields, and their linear equations of motion follow from the construction. Their two-point functions and commutators can be computed directly, and causality is automatic for fields associated with particles of integer spin. (In the massless case, one has to combine helicities $+h$ and $-h$.)

[1] In a famous panel discussion among P.A.M. Dirac, S. Ferrara, H. Kleinert, A. Martin, E.P. Wigner, C.N. Yang, and A. Zichichi, recorded in [42], it was consented that gauge theory does not refer to a “symmetry” of Nature. Its main role is the selection of renormalizable interactions that lie at the basis of the Standard Model of particles, without constituting a symmetry of the final quantum theory.

The spin-statistics theorem is the observation that in the case of half-integer spin, causality would conflict with the positivity of two-point functions viewed as inner products of one-particle state vectors, and the best one can achieve in this case is anti-locality—and this forces one to use a fermionic Fock space. Equal time (anti-) commutation relations are consequences, and they may not be "canonical" at all. Because anti-locality means maximal violation of causality, fermionic free fields are not observable fields. Their role is to create fermionic particle states, and to allow the construction of observable conserved currents. Their fate when an interaction is turned on has to be analyzed case by case.

For interactions, one adds an interaction Hamiltonian density, and appeals to time-dependent perturbation theory with a spacetime cutoff function $\chi(x)$ to begin with. In the end the adiabatic limit $\chi(x) \to 1$ has to be taken. This limit may be a tough problem, corresponding to IR problems popping up when one works without a cutoff function from the outset. The mathematically rigorous setup for perturbation theory with a cutoff, in which causality is under control, is known as "causal perturbation theory", or as the "Epstein–Glaser renormalization scheme (EG)" [12].

Common to both approaches is the resulting formula for the S-matrix in the vacuum sector

$$S = Te^{i \int d^4x \, L_{\mathrm{int}}(x)} \tag{1.1}$$

(to be properly interpreted), where L_{int} is a Wick polynomial in free fields. The S-matrix is computed perturbatively, by expansion of the exponential. In the EG approach, one replaces the coupling constant g in $L_{\mathrm{int}}(x)$ by a test function $g(x) = g\chi(x)$, and regards the resulting "local S-matrix" $S[g]$ as an operator-valued distribution with properties that can be axiomatized (and that may be established by a perturbative construction).

Worth mentioning is the absence of a "free Lagrangian L_0" in this formula,[2] which signals that its troublesome use for the mere purpose of "quantization" can (and should) be avoided altogether if possible. Indeed, the autonomous approach starts from the free fields, intrinsically defined in terms of the Wigner representations of the Poincaré group, and proceeds with (1.1).

In the standard approach, one has to postulate BRST invariance of the S-matrix, which otherwise cannot be unitary. This is a nontrivial issue, as $\delta_{\mathrm{BRST}} L_{\mathrm{int}}$ is (at lowest order) a total derivative which in general cannot be ignored because time-ordering does not commute with derivatives, see Sect. 2. This failure causes "obstructions" against BRST-invariance of the S-matrix, that have to be taken care of. In the form of a "principle of Perturbative Gauge Invariance (PGI)" [2, 34], the BRST method was developed to a mature tool to predict the structure of interactions, compatible with all sectors of the Standard Model. It is expanded in G. Scharf's book [34]. The technical (but not the conceptual!) aspects of this method are in fact a

[2] Path integral approaches use the "free Lagrangian" only for the formal definition of a (Euclidean) "Gaussian measure", the counterpart of the Fock space whose inner product is the free two-point function in real time.

precursor of the autonomous approach, which ultimately makes the BRST method obsolete altogether.

String-localized quantum fields. On the Wigner Fock space for helicity 1 (photons), there exist fields that one may oxymoronically call "gauge-invariant gauge potentials". The latter share formal properties with gauge potentials, but they do not contain unphysical modes, see Sect. 2.

This miracle is possible because one admits these potentials to have a weaker localization than usual: they may be "localized along a string". A "string" is a line or a narrow cone extending from x to spacelike infinity, and "localization" is understood in the sense of commutation relations. This feature is the origin of the name "string-localized quantum field theory (sQFT)". String-localized potentials also exist in the massive case: they have a better UV behaviour than local massive vector fields, because the weaker localization tames their vacuum fluctuations. This allows power-counting renormalizable interactions[3] that are otherwise forbidden.

The string-localized potentials are free fields, labelled by some auxiliary function $c(e)$ characterizing the string, as will be explained in Sect. 2. They are used to build interactions

$$L_{\rm int}(c) = gL_1(c) + \frac{g^2}{2}L_2(c) + \cdots$$

as a Wick polynomial (or possibly a power series) on the physical Hilbert space.

How sQFT determines interactions. (i) $L_{\rm int}(c)$ depends, through the string-localized potentials, on the auxiliary function $c(e)$. The point is that the S-matrix must not depend on this function ("string-independence"). At lowest order (in the coupling constant g), this implies that when c is varied by a function $\delta c(e)$, then $\delta_c L_1(c)$ must be a total derivative:

$$\delta_c L_1(x,c) = \partial_\mu Q_1^\mu(x,c,\delta c) \tag{1.2}$$

by virtue of the equations of motion of the free fields. We call (1.2) an "L–Q pair". The L–Q pair condition selects admissible first-order interactions, to begin with. They are typically cubic in the free fields. The first noticeable fact is that all known cubic interactions of the SM are among them, see Sect. 3 for their L–Q pairs.

Even more, in all cases of sQFT pertaining to SM interactions, there is also a point-localized interaction density $K_{\rm int} = gK_1 + \frac{g^2}{2}K_2 + \cdots$ to compare with. When the Hilbert space where $L_{\rm int}(c)$ is defined is suitably embedded into the possibly indefinite state space where $K_{\rm int}$ is defined, it holds

$$L_1(x,c) = K_1(x) + \partial_\mu V_1^\mu(x,c). \tag{1.3}$$

We call (1.3) an "L–V pair". It is the point of departure for an analysis [28, 31] to establish that $L_{\rm int}$ and $K_{\rm int}$ produce the same S-matrix, see below.

[3] We use the term "renormalizable" in the sense of power-counting throughout. The actual renormalization of sQFT has not been achieved so far, but the main result of [15] indicates that power-counting still is a valid criterion.

(ii) Once an initial L–Q pair (or L–V pair) is selected, it turns out that the higher orders of the S-matrix built with $L_1(c)$ will in general depend on the auxiliary function c: there arise "obstructions against string-independence" at higher order of the perturbative expansion. Their origin, as well as the strategy to cancel them is quite analogous to the failure and recuperation of BRST invariance in the PGI approach [34]: even if at first order (where there is no time-ordering) one has

$$\delta_c \int dx\, L_1(x,c) = \int dx\, \partial_\mu Q_1^\mu(x,c,\delta c) = 0$$

by virtue of (1.2), the second-order contribution

$$\delta_c \int dx\, dx'\; T[L_1(x,c)L_1(x',c)] = \int dx\, dx'\; T[\partial_\mu Q_1^\mu(x,c,\delta c)L_1(x',c) + (x \leftrightarrow x')] \tag{1.4}$$

will in general not vanish because the derivatives cannot be taken out of the time-ordered product. It must be cancelled by a second-order interaction L_2, which is only possible if the obstruction has a suitable form. In that case, it is said to be "resolved" by L_2, which is called the "induced" interaction. Restricting the analysis of the obstruction to tree level, it follows from Wick's theorem that if L_1 is cubic in the fields, then L_2 will be quartic. The second noticeable fact is that the induced interactions thus determined coincide with the known quartic interactions of the SM.

(iii) It may happen that obstructions are non-resolvable. A model with non-resolvable obstructions would have an S-matrix depending on the auxiliary string variable c without physical meaning, and must be abandoned as inconsistent with fundamental principles. The third noticeable fact is that the non-resolvable obstructions of several first-order interactions can cancel each other if all cross-channels are considered. We call this a "lock-key scenario": pieces of the SM will be inconsistent, while the whole becomes consistent. Examples are given in Sect. 3.

(iv) The method to compute induced interactions follows a recursive scheme, whose systematics has been elucidated in [31]. In principle, it may continue to arbitrary order. The fourth remarkable fact is that in all cases investigated so far, the recursion terminates at the power-counting bound for renormalizability—without renormalizability having been imposed.[4]

Interacting fields and observables. It is worth remembering that interaction densities $L_{\rm int}$ are just the field-theoretical tool to produce the corresponding S-matrix of interactions among (asymptotic) particles, by Formula (1.1). As long as the S-matrix computed with string-localized $L_{\rm int}(c)$ is independent of the string, it appears that one is done.

One should, however, object that a QFT is more than an S-matrix: it is defined by its interacting fields whose LSZ limit at $t \to \pm\infty$ is described by that S-matrix.

[4] The case of perturbative interactions involving massless particles of helicity 2 ("gravitons") is an exception, see Sect. 3.5: the lowest orders of induced interactions, that have been computed so far, reproduce the Einstein–Hilbert action which is known to be power-counting non-renormalizable already at first order, and there is no hint for a termination.

Especially *charged* fields are of interest because they create charged states from the vacuum. Interacting fields are perturbatively defined by Bogoliubov's formula

$$\Phi|_{L_{\text{int}}}(x) = \left(Te^{i\int L_{\text{int}}}\right)^* \left(T\left[\Phi(x)e^{i\int L_{\text{int}}}\right]\right), \tag{1.5}$$

to be expanded as a series of retarded integrals over free fields.

In gauge theory, interacting charged fields are not BRST invariant and are therefore not defined on the physical Hilbert space which is defined as the BRST quotient space (2.2). In the BRST approach, one must be content with an interacting quantum field theory lacking fields of major interest, that would generate charged states.

In sQFT, interacting fields can be constructed by the same Formula (1.5). Along with $L_{\text{int}}(c)$ being defined on the physical Hilbert space, so are the interacting fields. The point is, however, that a non-local interaction density in general will jeopardize locality of the perturbative interacting fields. Fortunately, there is a powerful tool [the "magic formula" (1.6)] in sQFT that secures at least string-locality. It proceeds by relating the interacting fields with interaction $L_{\text{int}}(c)$ to interacting "dressed" fields with a *local* interaction K_{int} as in the standard frameworks. The free "dressed" fields are much easier to compute than interacting fields, and their localization can be directly read off. The virtue of the magic formula is that (because K_{int} is local) the interacting dressed fields have the same localization as the free dressed fields, and this determines the localization of the interacting Hilbert space fields $\Phi|_{L_{\text{int}}(c)}$ even if $L_{\text{int}}(c)$ is string-localized.

Before we shall explain the working of the "magic formula", let us pause for a reflection about the meaning of "localization" in quantum field theory; and explain how sQFT allows to assess the localization of interacting fields.

Localization in quantum field theory is an algebraic feature—not a geometric one as the term might suggest (and of course not a matter of just labelling either). An operator is localized "somewhere" if it commutes with all observables at spacelike distance from "somewhere". Thus, localization refers to causality of the fields, rather than locality of the interaction.

A more precise formulation would be that localization is a *relative* algebraic property because it refers to the localization of observables. For example, in perturbation theory, interacting fields are represented as power series of retarded integrals over free fields. One may therefore not even expect them to be localized relative to the free fields, and Borchers [3] has shown that they cannot be so if there is nontrivial interaction.

But they can be localized relative to each other. This property selects interacting observables. Perturbed fields which fail to be localized relative to the interacting observables may still be present in the theory as non-observable fields, and they are welcome because they create states with non-local features [8]. One may call them "charged fields" in a generalized sense. Notice that anti-local fields (even free ones) violate causality in the strongest possible way. This is why only quadratic expressions like the Dirac *current* $j^\mu = :\overline{\psi}\gamma^\mu\psi:$ have a chance to qualify as observables, while the Dirac field ψ creates charged states.

In sQFT, string-localized *free* fields (notably the "gauge-invariant vector potentials" mentioned before) arise by integrating a free observable along a string emanating from x. These appear in $L_{\text{int}}(c)$, and define the perturbation theory. They are of little interest for the resulting *interacting* observables, but they leave their traces on the interacting charged fields.

The "magic formula" settles trouble with non-locality. A first reaction on sQFT is often the (in general justified) alarm that a non-local interaction will spoil locality of the interacting fields. We sketch here how locality is taken care of when there is an L–V pair available.

The all-orders analysis [31] of obstructions starting from an initial L–V pair (1.3) provides the prerequisites to investigate whether K_{int} and $L_{\text{int}}(c)$ produce the same S-matrix (at tree level at all orders). This indeed turns out to be the case for all interactions of the SM. Then, even if K_{int} may be non-renormalizable, e.g., because it contains couplings of massive vector bosons, the S-matrix $Te^{i\int dx\, K_{\text{int}}(x)}$ will be unitary and renormalizable—because $Te^{i\int dx\, L_{\text{int}}(x,c)}$ is. Conversely, the latter will be string-independent because the former is.

An offspring of the L–V analysis is the "magic formula"

$$\Phi\Big|_{L_{\text{int}}(c)}(x) = \Phi_{[g]}\Big|_{K_{\text{int}}}(x), \tag{1.6}$$

where the formula to compute the "dressed fields" $\Phi_{[g]}$ from the unperturbed field Φ, using the L–V pair, is given in [31]. This formula may transfer some of the degrees of freedom of the string-localized free fields onto the dressed fields. The latter may therefore depend on c and are manifestly relatively local or string-localized relative to the observable free fields (the latter can be characterized by the property that $\Phi^{\text{obs}}_{[g]}$ is local—e.g., if $\Phi^{\text{obs}}_{[g]} = \Phi^{\text{obs}}$ in the easiest case). By Borchers' result [3], this means that $\Phi_{[g]}$ are still not interacting—but their localization can be just read off. Then, because K_{int} is a local interaction density, standard results of causal perturbation theory assure that $\Phi_{[g]}|_{K_{\text{int}}}$ have the same *relative* localizations among each other as $\Phi_{[g]}$. Therefore, $\Phi|_{L_{\text{int}}}$ is relatively local resp. string-localized to the interacting observables. This is all one needs to know for the physical discrimination between "observable" or "charged" fields.

Some dressed fields, and consequently also the corresponding interacting fields are indeed local (and string-independent), and qualify as the local observables of the perturbative QFT. Some others (like the Dirac field of QED) become dynamically string-localized.

This bug is a feature: first, it is unproblematic because charged fields are not observables. Second, it is even to be expected in theories like QED, because charges can be measured and have effects at infinite spacelike distance by the Gauss Law. Notice that the Gauss Law does not violate causality, as the "creation of a charge" is not a physical process. The effects at spacelike infinity record the behaviour of the charges in the past.

The need to admit string-localized interacting fields was anticipated by an analysis (based on the axioms of algebraic quantum field theory referring exclusively to local observables) [8] of the causal support where charged superselection sectors (of massive theories) are distinguishable from the vacuum sector. The result can be interpreted as a statement on putative fields that would create such sectors from the vacuum, and conforms exactly with the weaker localization of the interacting fields in sQFT.

2 String-Localized Quantum Field Theory

We give here several complementary motivations for string-localized QFT. The first one starts from gauge theory and BRST and arises as an attempt to eliminate all obstructions against BRST invariance, as would appear in the PGI approach, from the start. Only the next two are truly in the autonomous spirit. They are enhancements towards applications in the SM of the original motivation [29] which was rooted in Algebraic QFT and "modular localization". This background will also be explained briefly below.

For simplicity, and in order to best exhibit the complementarity, we shall concentrate on the prototypical case of QED, with comments on other cases where appropriate.

String-localized quantum fields as the "reverse side of BRST". If one comes from gauge theory, one would start with a classical interaction density like

$$L_1(x) = A_\mu(x) j^\mu(x) \tag{2.1}$$

for QED. After covariant quantization, the field $A_\mu(x)$ is defined on a Fock space that contains states of negative norm square ("timelike photons") and states of zero norm ("longitudinal photons"). One eliminates these unphysical states by the Gupta-Bleuler method, or equivalently by the more general BRST method, which is the method of choice in non-abelian gauge theories. The latter extends the indefinite Fock space by another indefinite Fock space of "ghost" particles. On the extended Fock space $\mathcal{F}$ there is defined a nilpotent fermionic BRST operator Q with positive-semidefinite kernel. Its graded commutator $[Q, \cdot]_\mp$ defines the BRST variation $\delta_{\rm BRST} =: s$. Physical states (in QED: photons with only the two physical helicities, and no ghosts) are those annihilated by Q modulo states in the range of Q, which are of norm zero. In mathematical language, the physical Hilbert space is the quotient space ("cohomology")

$$\mathcal{H} = \mathrm{Ker}(Q)/\mathrm{Ran}(Q). \tag{2.2}$$

In order to be defined on this Hilbert space, observables must be BRST-invariant [$s(X) = 0$], including the S-matrix.

One finds that $s(j^\mu) = 0$ while

$$s(A_\mu) = \partial_\mu u, \tag{2.3}$$

where u is a ghost field. Consequently,

$$s(L_1) = \partial_\mu u j^\mu = \partial_\mu (u j^\mu) =: \partial_\mu P^\mu \tag{2.4}$$

is a total derivative. In particular, L_1 itself is not BRST invariant—only its classical action integral $\int d^4x\, L_1(x)$ is BRST invariant. But then the same is not automatically true for the perturbative S-matrix (1.1) because time-ordering does not commute with derivatives. For example,

$$s(T[L_1(x)L_1(x')]) = T[\partial_\mu P_1^\mu(x)L_1(x')] + (x \leftrightarrow x') \tag{2.5}$$

is in general not a derivative, in which case the integral will not vanish. This failure is called an "obstruction" against BRST invariance.

Fortunately, QED is particularly good-natured in that this obstruction actually does not occur, thanks to a Ward identity. (It is particularly bad-natured by its IR problems, though. The IR aspects of QED will be addressed separately in Sect. 3.1.) But the obstruction does occur in many other cases of interest, including the non-abelian case if L_1 is only the minimal interaction and the cubic part of the (self-) interaction. The quartic part of the self-interaction precisely cancels the obstruction. In [2, 34], the attitude is taken (and proven) that this cancellation mechanism is a justification for gauge invariant Lagrangians even if one does not assume gauge invariance from the start.

How would string-localized quantum fields enter the scene?

Because of (2.3) and (2.4), one seeks a field ϕ whose BRST variation would be

$$s(\phi) = -u. \tag{2.6}$$

Then, one could replace A_μ by the gauge-equivalent but BRST-invariant field $A_\mu + \partial_\mu\phi$, and L_1 by the BRST-invariant interaction density $(A_\mu + \partial_\mu\phi)j^\mu$.

Such a field ϕ is usually not "in the list" of BRST variations. But one can easily produce it: the string-localized integral

$$\phi(x,e) := \int_x^\infty dx'^\mu A_\mu(x') = \int_0^\infty ds\, e^\mu A_\mu(x+se) \tag{2.7}$$

does the job, with any string direction e.

One notices that (by virtue of the homogeneous Maxwell equations)

$$A_\mu(x,e) := A_\mu(x) + \partial_\mu\phi(x,e) = \int_0^\infty ds\, F_{\mu\nu}(x+se)e^\nu. \tag{2.8}$$

This representation of $A_\mu(x,e)$ makes its BRST-invariance more transparent: it is just a string integral over the BRST-invariant field strength. Equation (2.8) shows that $A_\mu(x,e)$ creates and annihilates only physical photon states with helicity ± 1, both manifestly (because F does so) and by counting degrees of freedom [namely, (2.8) implies $\partial^\mu A_\mu(x,e) = e^\mu A_\mu(x,e) = 0$].

Because of the distributional nature of quantum fields, one must smear the string direction e in (2.7), (2.8) with a suitable smooth function $c(e)$ and a suitable measure $d\sigma(e)$, thus defining $\phi(x,c)$ and $A_\mu(x,c)$.[5] If $\int d\sigma(e)\, c(e) = 1$ ("unit weight"), then

[5] In contrast to the smearing in x, we consider the smearing of the string variable as a label of the field $A(\cdot, c)$.

$$A_\mu(x,c) = A_\mu(x) + \partial_\mu \phi(x,c) \tag{2.9}$$

is still a potential for $F_{\mu\nu}$:

$$\partial_\mu A_\nu(x,c) - \partial_\nu A_\mu(x,c) = F_{\mu\nu}(x). \tag{2.10}$$

One may then work with the BRST-invariant interaction density

$$L_1(x,c) = A_\mu(x,c) j^\mu(x) = A_\mu(x) j^\mu(x) + \partial_\mu(\phi(x,c) j^\mu(x)). \tag{2.11}$$

Equation (2.11) is an L–V pair for QED. The S-matrix computed with $L_1(c)$ will be manifestly BRST-invariant. In the case of QED, one can establish that the derivative term does not alter the S-matrix. In more general cases, one may have to add higher-order interactions, as indicated after (1.4).

So, it seems that nothing is gained: instead of obstructions against BRST-invariance, there arise obstructions against string-independence, that must be resolvable. But on the contrary:

Because $L_{\rm int}(c) = gL_1(c) + \frac{g^2}{2} L_2(c) + \cdots$ is BRST-invariant, it is defined on the physical Hilbert space (2.2), and so will be the S-matrix and the interacting fields (1.5). One has a (perturbative) QFT on a Hilbert space without unphysical states from the outset, including charged fields that can generate charged states from the vacuum, whereas interacting charged fields in the BRST approach are not defined on the BRST Hilbert space. See the next paragraph.

The autonomous motivation for massless string-localized quantum fields. One can spare the detour through BRST by starting directly from the integral over F in (2.8)

$$A_\mu(x,c) := \int d\sigma(e)\, c(e) \int_0^\infty ds\, F_{\mu\nu}(x+se)e^\nu \tag{2.12}$$

as an *autonomous* definition of the string-localized potential satisfying (2.10), where F is defined on the Wigner Fock space. As in (2.11), the interaction is

$$L_1(x,c) = A_\mu(x,c) j^\mu(x), \tag{2.13}$$

with the distinction that there is no split (2.9) making it an L–V pair, because neither A nor ϕ are defined on the Wigner Fock space. The vital fact is that also without the split, $\delta_c(A_\mu(x,c))$ is still the derivative of a quantity defined on the Wigner Fock space, that we call $w(x,\delta c)$[6]:

$$\delta_c A_\mu(x,c) = \partial_\mu w(x,\delta c). \tag{2.14}$$

[6] Because c must have unit weight, a variation must be of the form $\delta c(e) = \partial^e_\kappa b^\kappa(e)$. The field $w(\delta c)$ is then computed by integrating ∂^e_κ by parts onto the s-integral in (2.8) where it produces a derivative ∂^x_κ, and using the homogeneous Maxwell equations. The resulting precise form of w is not of interest here. The field Q_1 contains a factor \$w\$, and may in more general models than QED also involve other string-localized fields.

Therefore, one has the L–Q pair of QED

$$\delta_c(L_1(x,c)) = \partial_\mu w(x,\delta c) j^\mu(x) = \partial_\mu(w(x,\delta c) j^\mu(x)) =: \partial_\mu Q_1^\mu(x,\delta c). \tag{2.15}$$

The next steps then follow the familiar scheme: perturbation theory with an interaction $L_1(c)$ will in general exhibit an obstruction against string-independence. In order to resolve the obstruction, one seeks a second-order interaction L_2 that makes

$$\delta_c\Big(Te^{i\int dx\left(gL_1+\frac{g^2}{2}L_2+\cdots\right)}\Big) = ig\int dx\,\delta_c(L_1(x,c)) + \frac{(ig)^2}{2}\Big[\iint dx\,dx'\,\delta_c\big(T[L_1(x,c)L_1(x',c)]\big) - i\int dx\,\delta_c(L_2(x,c))\Big] + \ldots \tag{2.16}$$

vanish. The first-order term vanishes thanks to (1.2). The condition that the second-order term vanishes at tree level is conveniently formulated as

$$O^{(2)}(x,x')\big|_{\text{tree}} \overset{!}{=} \delta_c(L_2(x,c))\cdot i\delta(x-x') + i\mathfrak{S}_2\partial_\mu^x Q_2^\mu(x,x',c,\delta c), \tag{2.17}$$

where $\mathfrak{S}_2$ is the symmetrization in two arguments, and

$$O^{(2)}(x,x') := 2\mathfrak{S}_2\Big(T[\partial_\mu Q_1^\mu(x,c,\delta c)L_1(x',c)] - \partial_\mu^x T[Q_1^\mu(x,c,\delta c)L_1(x',c)]\Big) \tag{2.18}$$

is the second-order obstruction that one would get from the first-order interaction alone. Equation (2.18) is analogous (with Q_1 in the place of a conserved current) to the violation of a Ward identity. By Wick's theorem, its tree-level part can be computed in terms of differences of time-ordered two-point functions (propagators)

$$O_\mu(\varphi(x),\chi(x')) := \langle T[\partial_\mu\varphi(x)\chi(x')]\rangle - \partial_\mu\langle T[\varphi(x)\chi(x')]\rangle \tag{2.19}$$

of the linear free fields of the model. The subtractions of explicit derivative terms bring the advantage that (2.19) and the numerical distributions in (2.18) automatically exhibit delta functions and string-integrated delta functions.

The existence of L_2 and Q_2 resolving the obstruction as in (2.17) is a nontrivial feature of a model. In particular, all string-localized delta functions must be part of the derivative term ∂Q_2. If non-resolvable obstructions arise, the model must be discarded, unless they can be cancelled (in the "lock-key scenario") by the inclusion of further first-order interactions. In general, one may have to proceed to higher orders.

The interacting fields $\Phi|_{L_{\text{int}}(c)}$ are computed with Bogoliubov's formula (1.5). In order to assess their localization with the help of the "magic formula" (1.6), one embeds the Wigner–Fock space into a suitable indefinite state space,[7] where there exists an L–V pair with a local interaction K_{int}. Then, $\Phi|_{L_{\text{int}}(c)}$ are defined on the

[7] In QED, the embedding can be made in such a way that the split (2.9) holds on a positive-definite subspace with a third degree of freedom besides the two polarizations of the photon. Upon interaction, this extra degree of freedom is transferred onto the interacting Dirac field—a feature whose importance has been stressed in [7]. (The photon continues to have two physical states.)

Hilbert space, but in general localized in the cone spanned by the strings e in the support of c.

In QED, the interacting Maxwell field turns out to be local, while the Dirac field is localized along the string, see Sect. 3.1. This feature resolves the conflict between the Gauss Law (the total electric charge equals the total electric flux at spacelike infinity), the Noether property [the total electric charge operator generates the $U(1)$ symmetry of the charged field], and locality (the electric flux at infinity should commute with the Dirac field at $x = 0$). There is no such conflict when the last property is dynamically invalidated. sQFT produces this effect intrinsically.

The autonomous motivation for massive string-localized quantum fields. A second beneficial feature of the Formula (2.8) arises with massive vector fields. The Proca field in the Wigner representation has short-distance scaling dimension 2: in its two-point function and propagator, $-\eta_{\mu\nu}$ is replaced by $-(\eta_{\mu\nu} + m^{-2}p_\mu p_\nu)$. The momentum-dependent term secures positivity but increases the UV scaling dimension, which is a measure for the size of vacuum fluctuations. This jeopardizes the use of the Proca field in perturbation theory: the minimal coupling to a current has dimension 5 beyond the renormalizability bound.

One way to deal with this is to just drop the term $m^{-2}p_\mu p_\nu$ which brings down the dimension to 1. The resulting field is a local "massive gauge field", whose two-point function violates positivity. One then needs BRST to return to a Hilbert space. BRST requires gauge invariance, but the mass violates gauge invariance. So the usual method of choice is to replace the massive gauge field by a massless one and invoke the Higgs mechanism to make it "behave as if it were massive".

Alternatively (see [34]), one may as well introduce a (local and positive) Stückelberg field which allows to adapt the BRST method without gauge invariance and ghosts. Also in this case, the BRST variation of L_1 is a derivative, causing obstructions of the S-matrix, which can be resolved only if the vector potential is coupled to a scalar field with the quartic interaction of the Higgs field.

The autonomous approach, in contrast, secures power-counting renormalizability by string-localization. The field strength of the Proca field of dimension 2 also has dimension 2 because the exterior derivative kills the "dangerous" momentum-dependent term in the two-point function. Then, because the subsequent integration decreases the dimension, the string-integral (2.12) with the massive Proca field strength (instead of the Maxwell field strength F) is a massive vector potential of dimension 1 on the Wigner–Fock space of the Proca field. When this field is coupled to a current, the renormalizability bound is respected.

For more on this treatment of the massive vector bosons of the electroweak interaction, including its ensuing prediction of chirality of the weak interactions and the shape of the Higgs self-coupling, see Sect. 3.4.

Modular localization. Localization in QFT is an algebraic feature: observables at spacelike separation must commute with each other. Usually, one constructs free fields $\varphi(x)$ whose commutator function is manifestly local. Upon perturbation with local scalar interactions, the relative localization is preserved.

In these approaches, the localization of an operator is encoded in the test function with which a local field is smeared. In contrast, "modular localization" is a construction of algebras of local observables that are distinguished by the relevant causal commutativity, without specific commutation relations when the separation is not spacelike. It proceeds with Modular Theory.

Modular Theory is an abstract theory about von Neumann algebras M on a Hilbert space with cyclic and separating vectors $\Omega \in \mathcal{H}$. From such a pair (M, Ω) it allows to extract its "modular data": a one-parameter group of automorphisms of M and an anti-linear conjugation j taking M to its commutant M'. These data have functorial properties which can be exploited for manifold applications of Modular Theory in QFT (by encoding locality in terms of commuting algebras), see [4].

"Modular localization" also starts from the Fock space over unitary Wigner representations. On this Hilbert space, one can specify von Neumann algebras M with the vacuum as a common cyclic and separating vector, in such a way that their modular data coincide with boost subgroups of the Poincaré group and PCT transformations [5]. (For the latter to be possible, one must combine the massless representations of helicities $+h$ and $-h$.)

Exploiting the correspondence between wedge regions of Minkowski spacetime and boost subgroups that preserve these regions, together with the functorial properties of modular data, one can *consistently define* algebras of local observables in wedge regions. Namely, the definition complies with the inclusion and commutation properties that are required for such an interpretation. In other words, the "localization" arises from Poincaré symmetry and modular data. Algebras of local observables in smaller regions can be defined as intersections of algebras for all wedge regions that contain the smaller regions.

Modular localization has not least opened the way to novel non-Lagrangian constructions of QFT models with interactions (in two dimensions so far), by deformations of von Neumann algebras [22].

The method applies as well to the "infinite spin representations" of the Wigner classification, for which local free fields do not exist [41]. Only in the last step of the agenda just outlined, the intersections of wedge algebras defining putative algebras of local observables in bounded regions, turn out to be trivial. The smallest intersections of wedge regions that admit nontrivial local observables are spacelike cones extending to infinity [23, 29].

This construction suggests that free fields for the infinite spin representations exist despite [41], provided they are allowed to be localized in cones. Indeed, the authors give concrete examples for such fields in [29]. They then notice that constructions like (2.12) are possible for every finite spin and helicity, and already observe their possible usefulness for perturbation theory, as outlined in the previous motivations.

The present paper intends to show that sQFT lives up to these expectations.

String-localized quantum field theory in action. In Sect. 3 we shall look at concrete realizations of sQFT in the Standard Model of elementary particles.

A main emphasis will be on the fact that interactions are *predicted* by the need to resolve obstructions, i.e., they are ultimately consequences of the underlying fundamental principles of quantum field theory: Positivity, covariance and locality, as explained in Sect. 1. The strategy is always to find first-order (cubic) interactions complying with the condition of string-independence (1.2), and study their obstructions to determine the higher interactions that make the S-matrix string-independent by "resolving the obstructions" as in (2.17).

Since string-independence is a *necessary* condition, it is legitimate to study the obstructions and their resolution only at tree level. Higher-order obstructions at tree level can always be reduced to expressions of a form generalizing (2.18), which in turn can be computed, with the help of Wick's theorem, in terms of propagators of the linear free fields of the model as in (2.19). The systematics at higher orders is fairly simple when one starts from L–Q pairs, and more involved with L–V pairs, see [31] for details.

The interactions determined by tree-level string-independence are then the starting point for the UV renormalization of the loop contributions, as in all other approaches. The power-counting bound being satisfied for $L_{\mathrm{int}}(c)$ is of course instrumental for this step. See Sect. 4.

The recurrent "lock-key scenario" has already been mentioned: The obstructions of certain first-order interactions cannot be resolved separately. One has to extend the model by other first-order interactions whose obstructions cancel each other. For example, the minimal interactions in non-abelian models require the cubic self-interactions of the vector bosons; and non-abelian self-interactions of massive vector bosons require interactions with a Higgs field [19]. Weak interactions of massive leptons require also Yukawa couplings to the Higgs field. Notice that these interactions are not added "in order to make massless particles massive", but they are required by consistency of interactions involving massive particles from the outset.

In this way, all the interactions otherwise known from gauge theory are accounted for. The resulting orthodox point of view of sQFT is that one should abandon "gauge invariance" altogether as a physical principle (not least because it exclusively refers to unobservable entities). However, we thank D. Buchholz for the comment that also imaginary numbers are "not real", yet they simplify life enormously. The same seems to be true for gauge theory methods as compared to sQFT (even if ghost fields may also be technically awkward). But the usefulness of gauge theory should not be mistaken as a "principle of Nature".

3 Examples and Achievements So Far

We give a brief overview of the manifold achievements of sQFT (some of them work in progress). Their relevance for the Standard Model of elementary particles will become apparent.

We shall present the relevant L–Q pairs and some of the resulting obstructions. The treatment of L–V pairs is technically more intricate, cf. [32], and will not be presented here, except for some results concerning dressed fields.

For detailed accounts, we refer to the cited literature.

3.1 QED

The initial L–Q and L–V pairs of QED were given in (2.15), (2.11). It turns out that obstructions of the S-matrix, as illustrated at second order in (2.18), vanish at all orders, and no induced higher-order interaction is needed.

Among the interacting fields, the Maxwell field tensor and the Dirac current remain local at all orders.

On the other hand, the dressed Dirac field is easily computed and turns out to be string-localized, so that also the interacting Dirac field is dynamically string-localized. The dressed Dirac field has the (superficially) simple form

$$\psi_{[g]}(x, c) = :e^{ig\phi(x,c)}: \psi(x), \tag{3.1}$$

where $\phi(x, c)$ was defined in (2.9). The function $c(e)$ must be supported within the unit sphere of a spacelike plane.[8] Quantities of the form (3.1) were previously considered by pioneers of QED [10, 21, 24] as *classical* expressions with the motive to quantize only gauge invariant quantities.

However, what looks like an innocent classical gauge transformation, has drastic consequences in quantum theory: the exponential of the infrared divergent quantum field ϕ turns the free Wigner Dirac field with a sharp mass into an infrafield which admits no mass eigenstates: its spectrum is only bounded from below. This spectral manifestation of the long-distance behaviour of QED was anticipated in a two-dimensional model [35], and later recognized as a consequence of the Gauss Law of QED in four dimensions [6].

Steinmann [36] also used quantum expressions of the form (3.1) as a starting point to formulate perturbation theory. His idea is very close to ours, with the distinction that he *chose* (3.1) (not least because it is gauge-invariant), while sQFT *derives* it.

The quantum field $\phi(c)$ in the exponent (3.1) is IR singular because of the integration over the massless vector potential involved in (2.7). But its exponential can be defined non-perturbatively with the help of an IR regularization [27]; and one finds that the string-smearing function $c(e)$ acquires a physical meaning: it describes the profile of the asymptotic electric flux density at infinity in the direction e, measured in states created from the vacuum by the field (3.1). Loosely speaking, it is the "shape of the photon cloud" of an electron.

This profile is not necessarily uniform in all directions (because physically it depends on the trajectory of the electron in the past), and it can be "engineered" by choosing the function $c(e)$ of unit weight.

States generated by the dressed Dirac field are orthogonal to the Fock space of the free fields because the coherent state created by the exponential factor has infinite particle number. In fact, the cloud function c defines a discrete superselection sector structure on an uncountably extended Hilbert space such that states with different

[8] The smearing with $c(e)$ is needed because of the distributional nature of quantum fields. It must be restricted to a spacelike plane in order not to spoil positivity.

c belong to different sectors. More precisely, for states with several electrons and positrons created by fields with different c, the sum (resp. difference) of their cloud functions is superselected.

While this is manifestly true for the free dressed fields, the uncountable superselection structure survives for the interacting Dirac fields, but seems to be "dynamically deformed" by the interaction $K_{\rm int}$ in (1.6) [27].

It was stressed in [7] that it is impossible to satisfy the Gauss Law with only two photon degrees of freedom, see Footnote 7. Indeed, unlike the potential $A_\mu(c)$, the field $\phi(c)$ is not defined on the photon Fock space. By restricting the support of $c(e)$ to a spacelike plane, its exponential remains positive but contributes one additional degree of freedom which becomes physical by turning the free Dirac field into an infrafield (3.1), and then describes its photon cloud. The interaction $K_{\rm int}$ then "hooks" at the dressing factor, so that the interacting field complies with the Gauss Law without "fictitious currents" (the failure of the free Maxwell equation in the Feynman gauge by a null field that, however, contributes to the interacting equations of motion). The field $\phi(c)$ itself is not part of the interacting theory because of its IR singularity.

The sQFT treatment of the infrared features of QED [27] encourages a new look at the "Infrared Triangle" (soft theorems—asymptotic symmetries—memory effect), see [37] for an overview. sQED provides conceptual (in the autonomous sense) elucidations of many of its features.

For example, sQED allows to compute quantum expectation values of the Maxwell field in states created by dressed charged fields, and study their asymptotic behaviour at spacelike ($\mathfrak{i}^0$), null ($\mathfrak{I}^\pm$), and timelike ($\mathfrak{i}^\pm$) infinity.

The behaviour along $\mathfrak{I}^\pm$ can be regarded as a pair of three-dimensional QFTs with their own intrinsic algebraic structure and dynamics, arising as asymptotic limits from the bulk, and capturing the helicity-one version of the gravitational "memory effect". The "matching conditions" between the past rim of future null infinity and the future rim of past null infinity can be traced by following the continuous interpolation from $\mathfrak{I}^-$ to $\mathfrak{I}^+$ across spacelike infinity $\mathfrak{i}^0$.

When one views the matching conditions as conservation laws [37], the corresponding infinitely many electric and magnetic conserved charge operators Q_ε and $\widetilde{Q}_\varepsilon$ can be computed, along with the "large gauge transformations" that they would generate in the Feynman gauge. In fact, the latter act also on the additional degree of freedom in the field $\phi(c)$, emphasized above, and consequently on the dressed field and the state it creates from the vacuum. Thus, while they leave the Maxwell field invariant, they transform its asymptotic expectation values.

The connection with soft photon scattering is manifest in sQED in the dynamical deformation of the superselection structure [27] caused by the proper (positivity-preserving) treatment of the IR divergent field $\phi(c)$ in the exponent in (3.1).

3.2 Yang–Mills

A cubic L–Q pair involving several massless vector potentials $A^a_\mu(c)$ is necessarily of the form

$$L_1(c) = -\frac{1}{2}\sum_{abc} f_{abc}F^{a\mu\nu}A^b_\mu(c)A^c_\nu(c), \quad Q^\mu_1(c) = -\sum_{abc} f_{abc}F^{a\mu\nu}w^b(\delta c)A^c_\nu(c). \tag{3.2}$$

A most general ansatz shows that the numerical coefficients f_{abc} must be totally antisymmetric (string-independence at first order) and must satisfy the Jacobi identity (second order) [16]. Thus, they are necessarily the structure constants of an (unspecified) reductive Lie algebra $\mathfrak{g}$. (The same result was also found within the PGI approach, see [2].)

If $\mathfrak{g}$ is non-abelian, there is a second-order obstruction that can be resolved by the induced quartic interaction

$$L_2(c) = -\frac{1}{2}\sum_{abcde} f_{abe}f_{cde}A^a_\mu(c)A^b_\nu(c)A^{c\mu}(c)A^{d\nu}(c),$$

and higher-order interactions do not appear. $gL_1(c) + \frac{g^2}{2}L_2(c)$ is precisely the string-localized version of the Yang–Mills interaction.

3.3 QCD

One may add to (3.2) an L–Q pair of minimal quark-gluon interactions [copies of (2.15) with colored currents $j^\mu_a = \overline{\psi}\gamma^\mu\tau_a\psi$]. The minimal interaction, if taken separately, has a non-resolvable obstruction at second order:

$$-\sum_{abc} f_{abc}w^a(x,\delta c)A^b_\mu(x,c)j^\mu_c(x)\cdot\delta(x-x'), \tag{3.3}$$

which arises from the violation of the Ward identity for the non-abelian currents. It is resolved by the same obstruction with the opposite sign appearing in the cross contributions when the self-interaction (3.2) is included [32]. In particular, there is no induced interaction in the quark sector. This is the first appearance of the "lock-key scenario" mentioned in Sect. 2.

The computation of the dressed quark field is much more difficult than in QED, because of the non-abelianness. We do not have a closed expression like (3.1), but the first three orders [32] indicate the onset of a path-ordered exponential

$$\psi_{[g]}(x,c) = {:}Pe^{ig\phi(x,c)}{:}\cdot\psi(x). \tag{3.4}$$

Recall that $\phi(e) = \sum_a I_e(A^a)\tau_a$ is a Lie-algebra-valued line integral over A. The path-ordering orders the Lie-algebra generators, while the field operators are Wick-ordered. Path-ordering is well defined for a sharp string ($c(e)$ supported in a single direction) and the resulting Wilson operator makes (3.4) invariant under *classical* gauge transformations that are trivial at infinity. For smooth smearing functions $c(e)$ (which are needed because of the distributional nature of quantum fields), "path-ordering" of the exponential of smeared line integrals is not defined in an obvious way, but (3.4) still enjoys gauge-invariance (to third order). To give a flavour of the result [32]:

$$\psi_{[g]}(x,c) = \,{:}\!\exp i\left[g\phi(c) + \frac{g^2}{2}I_c(i[\phi(c), 2A + \partial\phi(c)]) + \frac{g^3}{6}\phi^{(3)}(c) + \cdots\right](x){:}\cdot\psi(x), \tag{3.5}$$

where $A = A(c) - \partial\phi(c)$ is the Lie-algebra-valued local potential in Feynman gauge,[9] $i[\cdot,\cdot]$ is the Lie algebra commutator, and $I_c(Y)$ is short-hand [appearing already in (2.12), where we had $\phi(x,c) = (I_c(A))(x)$] for the string integral operation on a vector field

$$(I_c(Y))(x) := \int d\sigma(e)\, c(e) \int_0^\infty ds\, e_\mu Y^\mu(x + se).$$

The third-order term in (3.5) is

$$\phi^{(3)} = 3I_c(i[I_c(i[\phi, 2A + \partial\phi]), A + \partial\phi]) + I_c(i[\phi, i[\phi, 3A + \partial\phi]]) - \frac{3}{2}i[\phi, I_c(i[\phi, 2A + \partial\phi])].$$

For sharp strings, the nested string integrations of commutators in the exponent take care of the path-ordering along the string.

As in QED, the string-integrated field $\phi = I_c(A)$ is infrared divergent, and its appearance in the exponent requires a non-perturbative definition. We speculate that analytic confinement criteria such as [40] or [13] could then be established already at the level of the dressed quark field (3.4); but for this one would need a closed formula at all orders which is presently not available.

3.4 Electroweak Interactions

Turning to weak interactions, one may as well generalize the *L*–*Q* pair (3.2) to *massive* vector bosons, including self-coupling terms proportional to the masses [18, 19]. One may again add *L*–*Q* pairs for minimal couplings to vector or axial currents $J^\mu_a = c_V j^\mu_a + c_A j^{5\mu}_a$. For massive fermions (because the axial current is not conserved), one also needs couplings to scalar and pseudoscalar currents with coefficients determined at second order.

The fermionic sector with the exact field content and masses of the SM (including the Higgs field) has been analyzed, and its details are reported in [18]. The main emphasis is that the maximal chirality ($c_A = \pm c_V$) of the weak coupling of the *W*-bosons is a necessary condition for the resolution of obstructions, which needs both the minimal interactions and the self-interactions of the vector bosons.

9 The expression (3.5) is not manifestly positive. But it *is* positive, because it was computed with the help of an *L*–*V* pair on a positive-definite subspace of the Feynman gauge Fock space, see Footnote 7.

Chirality is of course well-known as an empirical fact and used as a constitutive feature of the GSW model, but has no a priori explanation from gauge theory. Also the empirically known mass-dependence of the Yukawa couplings follows as a consequence of string-independence at first and second order.

The bosonic sector is presently under investigation [19], including the self-interaction of massive and massless vector bosons and their couplings to the Higgs boson. The most important feature is that the latter are necessary in order to resolve non-resolvable obstructions of the self-interaction. Moreover, the masses m_W and m_Z as the only empirical input fix, as a consequence of string-independence, the Z and photon couplings to W, and the coupling strengths of Z and W to the Higgs boson. There arise non-resolvable obstructions at third order in the sectors with one or three Higgs fields. They can be cancelled if one adds Higgs self-interactions ℓH^3 to L_1, and $\ell' H^4$ to L_2. The values ℓ and ℓ' determined by string-independence are precisely the same as one also would get from the shifted-symmetric double-well potential

$$\frac{1}{2}\lambda\left[(v+H)^2 - v^2\right]^2 \quad \text{with} \quad gv = 2m_W, \quad \lambda = m_H^2/4v^2$$

of the GSW model. (Details of this latter analysis are almost identical with the abelian model as reported in [27].)

Notice that sQFT predicts a Higgs particle along with its self-interaction [19]—but without a "Higgs mechanism" because the vector bosons are massive from the start.

Let us point out an interesting pattern of cancellation of obstructions, that determines the Lie algebra. It arises most cleanly in the simpler Higgs–Kibble model, anticipating a similar one in the full electroweak theory. The model describes massless fermions and vector bosons of equal mass and no photon [32].

The L–Q pair for the self-coupling is (with summation over abc understood)

$$\begin{aligned} L_{1,\text{self}}(c) =& \frac{1}{2} f_{abc}\left(F^{a\mu\nu} A^b_\nu(c) - m^2 \phi^a(c) B^{b\mu}\right) A^c_\mu(c), \\ Q^\mu_{1,\text{self}}(c) =& f_{abc}\left(F^{a\mu\nu} A^b_\nu(c) - \frac{1}{2} m^2 \phi^a(c) B^{b\mu}\right) w^c(\delta c) \end{aligned}$$

where $B_\mu = A_\mu(c) - \partial\phi(c)$ is the local Proca field.

The self-coupling has a non-resolvable obstruction at second order of the form

$$m^2 \cdot f_{abe} f_{cde} \cdot w^a(\delta c) A^b_\mu(c) \phi^c(c) B^{d\mu}. \tag{3.6}$$

This obstruction can be cancelled by another non-resolvable obstruction arising from the coupling to a scalar field (the Higgs field of mass m_H) with L–Q pair (summation over a understood)

$$\begin{aligned} L_{1,\text{Higgs}}(c) =& \mu \cdot \left(A^a_\mu(c) B^{a\mu} H + A^a_\mu(c)\phi^a(c)\partial^\mu H - \frac{1}{2} m_H^2 \phi^a(c)\phi^a(c) H\right), \\ Q_{1,\text{Higgs}}(c) =& \mu \cdot w^a(\delta c)(B^{a\mu} H + \phi^a(c)\partial^\mu H). \end{aligned}$$

Its non-resolvable obstruction has the form

$$-4\mu^2 \cdot (\delta_{ab}\delta_{cd} - \delta_{ad}\delta_{cb}) \cdot w^a(\delta c)A^b_\mu(c)\phi^c(c)B^{d\mu}. \qquad (3.7)$$

The cancellation of (3.6) against (3.7) requires to match the self-coupling coefficients f_{abc} with the Higgs coupling μ via $m^2 \cdot f_{abe}f_{cde} = 4\mu^2 \cdot (\delta_{ab}\delta_{cd} - \delta_{ad}\delta_{cb})$. This determines the structure constants $f_{abc} = \varepsilon_{abc}$ of $\mathfrak{su}(2)$ and $\mu = \frac{1}{2}m$. (The overall relative factor $\frac{1}{2}$ as compared to the abelian model [27] corresponds to the normalization convention of the generators $\frac{1}{2}\sigma_a$ of $\mathfrak{su}(2)$.) Other Lie algebras are presumably possible when nontrivial mass patterns or several Higgs fields are admitted [19].

3.5 Helicity 2

In [17], three of us have pursued an analysis of interactions of massless particles of helicity ± 2 ("gravitons") with massive or massless matter of spin 0, $\frac{1}{2}$ or 1. The latter do not need sQFT, but the former do, because there is no local covariant Hilbert space field to serve as a "metric deviation" $h_{\mu\nu}(x)$. There is, however, a string-localized field $h_{\mu\nu}(x, c)$, built from the "linearized Riemann tensor field" on the Fock space over the unitary Wigner representations of helicity ± 2, by a formula analogous to (2.12) (with two string integrations).

The analysis of possible cubic L–Q pairs results in unique couplings $\frac{1}{2}h_{\mu\nu}(c)\Theta^{\mu\nu}_{\text{matter}}$ to the conserved matter stress–energy tensors, and in a unique (up to total derivatives) cubic self-coupling. The matter couplings separately have non-resolvable obstructions, which are all cancelled in the cross-channels with the self-coupling. This is another appearance of the lock-key scenario. More remarkably, the induced interactions (computed only at second order) coincide with the classical expansions of the Einstein–Hilbert action and of the standard generally covariant matter couplings, upon substitution of the classical metric deviation field $h_{\mu\nu}$ by the quantum field $h_{\mu\nu}(c)$.

General covariance was not assumed. It arises instead as a consequence of the quantum principles underlying sQFT—a feature that we called "quantum general covariance".

Clearly, the expansion is not power-counting renormalizable. Yet, the strong constraints underlying the resolution of obstructions may even raise the hope that the renormalized sQFT of gravity has no free parameters besides the coupling constant.

A similar result was obtained in the PGI setting [11]: consistency with BRST, which in turn is needed to secure positivity, is secured by the expansion of the Einstein–Hilbert self-interaction in all orders.

3.6 Beyond the Standard Model?

There are many proposals that "dark matter" might consist of particles of spin or helicity beyond 2. We want to discuss what could be possible with sQFT in this respect.

Recall that by the Weinberg–Witten theorem, particles of helicity 2 or higher do not admit local and covariant stress–energy tensor fields on a Hilbert space. This is often taken as an argument that massless matter of higher helicity cannot couple gravitationally, see a discussion in [30]. Stress–energy tensors of massive matter of higher spin exist but have increasing UV scaling dimension, leading to "more and more non-renormalizable" interactions.

On the contrary, there exist string-localized stress-energy tensors $\Theta^{\mu\nu}_{\rm matter}(c)$ of UV dimension 4 for every integer spin and helicity [25]. The self-coupling of the helicity-2 field mentioned in Sect. 3.5 is a small modification of $\frac{1}{2}h_{\mu\nu}(c)\Theta^{\mu\nu}_{h=2}(c)$ (which would be part of an L–Q pair only if $\Theta_{h=2}$ were string-independent). This observation raises the hope that sQFT might also admit L–Q pairs for graviton couplings to massless matter of helicity beyond 2. Unfortunately, the systematic search turned out no candidates beyond the known ones with $h = 0, 1, 2$. The analysis of L–Q pairs coupling gravitons to particles with half-integer helicity and to massive higher-spin particles remains to be done.

One may also wonder whether sQFT allows new interactions among known particles. One such proposal is [1]. The authors observed that with an interaction L' involving the Z_μ field multiplying two dressing factors $\exp ig\phi(e)$ and $\exp -ig\phi(e')$ with $e \neq e'$, one can produce the decay of a Z-boson into two photons, otherwise forbidden by the Landau–Yang theorem. (The latter uses Lorentz invariance which is manifestly broken by the choice of the strings.)

The proposed interaction term L' is not part of an L–Q or L–V pair, so the resulting scattering matrix has no reason to be string-independent. This may not even be intended by the authors; but what weighs heavier is that as a consequence, there is no "magic formula" which would ensure any kind of localization for quantum fields subjected to an interaction L', along the lines discussed in Sect. 1.[10]

We do not know what is possibly "allowed" as string-localized interactions beyond L–Q or L–V pairs and higher-order interactions induced from them, and how much flexibility is gained for interactions beyond the SM. For example, L–Q pairs that are quartic in the free fields may have slipped our attention. But we strongly feel that one cannot choose any interaction merely because it seems convenient. Locality and positivity are precious goods that must be taken care of.

4 Conclusion and Outlook

String-localized quantum field theory (sQFT) is a new framework to address the relation between particles and quantum fields. For free fields, it is just "more flexible" than local QFT. This flexibility can be used to overcome problems with Hilbert space positivity due to canonical quantization of massless vector fields, and problems with renormalizability due to the increased UV scaling dimension of vacuum fluctuations of massive vector fields [25].

[10] The authors use an ad hoc IR and UV regularization which in fact violates positivity. Smearing the strings with $c(e)$ would solve the UV issue. The positivity-preserving IR regularization used in [26] produces the superselection structure mentioned in Sect. 3.1, and would actually make the amplitude in [1] vanish whenever $c \neq c'$.

sQFT deploys its power when applied to perturbation theory: it strongly constrains admissible interactions by fundamental quantum principles (notably Hilbert space). In fact, sQFT can be understood as a scheme to predict the interactions of the SM in an "autonomous" (purely quantum) way.

A prominent lesson that sQFT teaches us is that perturbed quantum fields, except the observables of the model, must not be expected to be local (or anti-local). Their dynamical string-localization is a physical feature (e.g., related to the Gauss Law of QED), and it still allows "sufficient commutativity" to conceive a scattering theory in terms of correlations at asymptotic times. However, the prevalent methods of scattering theory have to be properly adjusted.

The dressed Dirac and quark fields (3.1) and (3.4), that are (classically) gauge-invariant, are instances of an issue that calls out for a deeper understanding: given that sQFT a priori "knows nothing" about gauge theory, why do (classically) gauge-invariant string-localized fields play a distinguished role in sQFT, and why seem the observables of sQFT to correspond to (classically) gauge-invariant local fields? Is perhaps "gauge-invariant and local" just another characterization of "string-independent"?

An interesting recent argument by Rivat [33] proposes another reason for "gauge invariance": it explains it as a consequence of the need to have a Lorentz invariant action $\int d^4x\, L_{\rm int}(x)$. This looks unrelated to sQFT, but there is a close connection. Rivat starts from the well-known facts that on the Wigner Fock space of the photon, a Lorentz covariant vector potential does not exist [38]; but one can construct a vector potential $A^{\rm W}_\mu$ which is Lorentz covariant "up to an operator-valued gauge transformation":

$$U(\Lambda)A^{\rm W}_\mu(x)U(\Lambda)^* = \Lambda^\nu{}_\mu\big(A^{\rm W}_\nu(\Lambda x) + \partial_\nu\Gamma(\Lambda,x)\big), \tag{4.1}$$

where Γ arises from the "pseudo-translations" in the little group $E(2)$ (the stabilizer subgroup of the Lorentz group of an arbitrary massless reference momentum vector) in the Wigner decomposition of Λ. A transformation law of the same form as (4.1) holds for $A_\mu(c)$, where Γ arises instead from the Lorentz transformation of the strings. Consequently, interactions like both $A^{\rm W}_\mu j^\mu$ and $A_\mu(c)j^\mu$ change by a total derivative under Lorentz transformations. The main difference between $A^{\rm W}_\mu$ and $A_\mu(c)$ is the well-controlled localization of the latter.

The focus in this paper was on the relevant analysis to *determine* interactions, which proceeds at tree-level. The development of an efficient UV renormalization program is the main open agenda of sQFT. An adaptation of the EG program [12] to the string-localized context might be most promising.

A general strategy can be outlined already. With a fixed smearing function $c(e)$, string-localized two-point functions enjoy the Hadamard property [14], which allows to formulate a Wick expansion. Replacing two-point functions by kinematic propagators, one obtains an unrenormalized time-ordered Wick expansion, which has to be extended to the singular set where it is not well-defined. Thanks to the Hadamard property, this set is not bigger than in the local case [14, 15], and in particular, the usual power-counting method applies, limiting the scaling degree of the renormalizations. But there may be more ambiguities in the extensions because of the possibility of string-integrated delta functions. Their classification seems to be the main difficulty, where a "factorization rule" for string-localized S-matrized would be instrumental. Here, ideas from "string-chopping" [9] are expected to be helpful. At the same time, the increase of ambiguities will be (partly) counter-balanced by the condition that string-independence (valid at tree level) must be preserved.

On the other hand, the equivalence results at tree-level, presented in this paper and [31], let us surmise that sQFT will not produce other results than local QFT with its "negative probabilities plus BRST". One may therefore do the actual loop calculations in the usual framework, while the message of sQFT is conceptual: The interactions of the SM can be understood without gauge fields and ghosts, and are rather direct consequences of fundamental principles of quantum field theory: Hilbert space positivity, covariance and locality.

Acknowledgements We thank J. Mund and J. Yngvason who contributed many substantial ideas (some of them unpublished) in earlier stages of the line of research reported here. We thank D. Buchholz for interesting and constructive comments. KHR thanks his former MSc and BSc students, M. Chantreau, I. Hemprich, T. Joswig, L. Kersten, N. Münch, M. Pabst and F. Tippner, who with their thesis works have been involved in various explicit computations in some of the models of Sect. 3 and [32]. CG was funded by the National Science Center of Poland under the Grant UMO-2019/35/B/ST1/01651.

Author Contributions All authors have contributed to various portions of the research results reported in the manuscript. KHR has first edited the text. All authors have reviewed and improved the draft.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Asorey, M., Balachandran, A.P., Arshad Momen, M., Qureshi, B.: Violation of the Landau–Yang theorem from infrared Lorentz symmetry breaking. JHEP **10**, 028 (2023)
2. Aste, A., Scharf, G.: Non-abelian gauge theories as a consequence of perturbative quantum gauge invariance. Int. J. Mod. Phys. A **14**, 3421–3432 (1999)
3. Borchers, H.-J.: Über die Mannigfaltigkeit der interpolierenden Felder zu einer kausalen S-Matrix (in German). Nuovo Cim. **15**, 784–794 (1960)
4. Borchers, H.-J.: On revolutionizing quantum field theory with Tomita's modular theory. J. Math. Phys. **41**, 3604–3673 (2000)

5. Brunetti, R., Guido, D., Longo, R.: Modular localization and Wigner particles. Rev. Math. Phys. **14**, 759–785 (2002)
6. Buchholz, D.: Gauss' Law and the infraparticle problem. Phys. Lett. B **174**, 331–334 (1986)
7. Buchholz, D., Ciolli, F., Ruzzi, G., Vasselli, E.: On string-localized potentials and gauge fields. Lett. Math. Phys. **109**, 2601–2610 (2019)
8. Buchholz, D., Fredenhagen, K.: Locality and the structure of particle states. Commun. Math. Phys. **84**, 1–54 (1982)
9. Cardoso, L.T., Mund, J., Várilly, J.C.: String chopping and time-ordered products of linear string-localized quantum fields. Math. Phys. Anal. Geom. **21**, 3 (2018)
10. Dirac, P.A.M.: Gauge-invariant formulation of Quantum Electrodynamics. Can. J. Phys. **33**, 650–660 (1955)
11. Dütsch, M.: Proof of perturbative gauge invariance for tree diagrams to all orders. Ann. Phys. (Leipzig) **14**, 438–461 (2005)
12. Epstein, H., Glaser, V.: The role of locality in perturbation theory. Ann. Inst. Henri Poincaré A **19**, 211–295 (1973)
13. Fredenhagen, K., Marcu, M.: Confinement criterion for QCD with dynamical quarks. Phys. Rev. Lett. **56**, 223–224 (1986)
14. Gass, C.: Constructive aspects of string-localized quantum field theory. PhD Thesis, Universität Göttingen (2022)
15. Gass, C.: Renormalization in string-localized field theories: a microlocal analysis. Ann. Inst. Henri Poincaré **23**, 3493–3523 (2022)
16. Gass, C., Gracia-Bondía, J.M., Mund, J.: Revisiting the Okubo–Marshak argument. Symmetry **13**, 1645 (2021). https://doi.org/10.3390/sym13091645
17. Gass, C., Gracia-Bondía, J.M., Rehren, K.-H.: Quantum general covariance. Class. Quantum Gravity **40**, 195016 (2023)
18. Gracia-Bondía, J.M., Mund, J., Várilly, J.C.: The chirality theorem. Ann. Inst. Henri Poincaré **19**, 843–874 (2018). https://doi.org/10.1007/s00023-017-0637-3
19. Gracia-Bondía, J.M., Rehren, K.-H., Várilly, J.C.: The full electroweak interaction: an autonomous account (in preparation)
20. Jordan, P.: Der gegenwärtige Stand der Quantenelektrodynamik (in German). In: Talks and discussions of the theoretical–physical conference in Kharkov (19–25 May 1929). Phys. Z. **XXX**, 700–712 (1929)
21. Jordan, P.: Zur Quantenelektrodynamik (in German). Z. Phys. **95**, 202–209 (1935)
22. Lechner, G.: Construction of quantum field theories with factorizing S-matrices. Commun. Math. Phys. **277**, 821–860 (2008)
23. Longo, R., Morinelli, V., Rehren, K.-H.: Where infinite spin particles are localizable. Commun. Math. Phys. **345**, 587–614 (2016)
24. Mandelstam, S.: Quantum electrodynamics without potentials. Ann. Phys. **19**, 1–24 (1962)
25. Mund, J., Rehren, K.-H., Schroer, B.: Helicity decoupling in the massless limit of massive tensor fields. Nucl. Phys. B **924**, 699–727 (2017)
26. Mund, J., Rehren, K.-H., Schroer, B.: Gauss' Law and string-localized quantum field theory. JHEP **01**, 001 (2020). https://doi.org/10.1007/JHEP01(2020)001
27. Mund, J., Rehren, K.-H., Schroer, B.: Infraparticle quantum fields and the formation of photon clouds. JHEP **04**, 083 (2022). https://doi.org/10.1007/JHEP04(2022)083
28. Mund, J., Rehren, K.-H., Schroer, B.: How the Higgs potential got its shape. Nucl. Phys. B **987**, 116109 (2023)
29. Mund, J., Schroer, B., Yngvason, J.: String-localized quantum fields and modular localization. Commun. Math. Phys. **268**, 621–627 (2006)
30. Porrati, M.: Old and new No Go theorems on interacting massless particles in flat space. In: 17th International Seminar on High Energy Physics (Quarks 2012), 2012. arXiv:1209.4876
31. Rehren, K.-H.: On the effect of derivative interactions in quantum field theory. arXiv:2405.09168
32. Rehren, K.-H.: Nonabelian models in sQFT. arXiv:2405.09370
33. Rivat, S.: Wait, why gauge? Br. J. Philos. Sci. (2024). https://doi.org/10.1086/727736
34. Scharf, G.: Gauge Field Theories: Spin One and Spin Two. Dover Publications, Mineola (2016)
35. Schroer, B.: Infraparticles in quantum field theory. Fortsch. Phys. **11**, 1–32 (1963)
36. Steinmann, O.: Perturbative QED in terms of gauge invariant fields. Ann. Phys. **157**, 232–254 (1984)

37. Strominger, A.: Lectures on the Infrared Structure of Gravity and Gauge Theory. Princeton University Press, Princeton (2018)
38. Weinberg, S.: Feynman rules for any spin. Phys. Rev. **138**, 1318–1332 (1964)
39. Wigner, E.P.: On unitary representations of the inhomogeneous Lorentz group. Ann. Math. **40**, 149–204 (1939)
40. Wilson, K.: Confinement of quarks. Phys. Rev. D **10**, 2445–2459 (1974)
41. Yngvason, J.: Zero-mass infinite spin representations of the Poincaré group and quantum field theory. Commun. Math. Phys. **18**, 195–203 (1970)
42. Zichichi, A. (ed.): Panel Discussion, Recorded in: "Gauge Interactions: Theory and Experiment, pp. 725–740. Plenum, New York (1984)

Authors and Affiliations

K.-H. Rehren[1] · L. T. Cardoso[2] · C. Gass[3] · J. M. Gracia-Bondía[4] · B. Schroer[5] · J. C. Várilly[6]

✉ K.-H. Rehren
krehren@uni-goettingen.de

L. T. Cardoso
lucas.cardoso@ufsm.br

C. Gass
christian.gass@fuw.edu.pl

J. M. Gracia-Bondía
jose.gracia@ucr.ac.cr

B. Schroer
schroer@zedat.fu-berlin.de

J. C. Várilly
joseph.varilly@ucr.ac.cr

1 Institut für Theoretische Physik, Georg-August-Universität Göttingen, 37077 Göttingen, Germany

2 Coordenadoria Acadêmica, Universidade Federal de Santa Maria, Cachoeira do Sul, RS 96503-205, Brazil

3 Department of Mathematical Methods in Physics, University of Warsaw, 02-093 Warszawa, Poland

4 Laboratorio de Física Teérica y Computacional, Universidad de Costa Rica, San Pedro 11501, Costa Rica

5 Institut für Theoretische Physik, Freie Universität Berlin, 14195 Berlin, Germany

6 Escuela de Matemática, Universidad de Costa Rica, San Pedro 11501, Costa Rica