The Holst spin foam model via cubulations

Aristide Baratin[1]**, Cecilia Flori**[2] **and Thomas Thiemann**[3]

[1] Max Planck Institute for Gravitational Physics, Albert-Einstein-Institute,
Am Mühlenberg 1, 14476 Potsdam, Germany
[2] Perimeter Institute for Theoretical Physics, 31 Caroline Street N, Waterloo,
ON N2L 2Y5, Canada
[3] Institute for Theoretical Physics III, University of Erlangen-Nürnberg,
Staudtstrasse 7, 91058 Erlangen, Germany
E-mail: aristide.baratin@aei.mpg.de, cflori@perimeterinstitute.ca
and thiemann@aei.mpg.de

Abstract. Spin foam models are an attempt at a covariant or path integral formulation of canonical loop quantum gravity. The construction of such models usually relies on the Plebanski formulation of general relativity as a constrained BF theory and is based on the discretization of the action on a simplicial triangulation, which may be viewed as an ultraviolet regulator. The triangulation dependence can be removed by means of group field theory techniques, which allows one to sum over all triangulations. The main tasks for these models are the correct quantum implementation of the Plebanski constraints, the existence of a semiclassical sector implementing additional 'Regge-like' constraints arising from simplicial triangulations and the definition of the physical inner product of loop quantum gravity via group field theory. Here we propose a new approach to tackle these issues stemming directly from the *Holst action* for general relativity, which is also a proper starting point for canonical loop quantum gravity. The discretization is performed by means of a 'cubulation' of the manifold rather than a triangulation. We give a direct interpretation of the resulting spin foam model as a generating functional for the *n*-point functions on the physical Hilbert space at finite regulator. This paper focuses on ideas and tasks to be performed before the model can be taken seriously. However, our analysis reveals some interesting features of this model: firstly, the structure of its amplitudes differs

from the standard spin foam models. Secondly, the tetrad n-point functions admit a 'Wick-like' structure. Thirdly, the restriction to simple representations does not automatically occur—unless one makes use of the time gauge, just as in the classical theory.

Contents

1. Introduction

Spin foam models (SFM) [1] are an attempt at a covariant or path integral formulation of canonical loop quantum gravity (LQG) [2–4]. In their current formulation (e.g [5–9]), SFM exploit the Plebanski formulation [10] of pure general relativity (GR) as a constrained BF theory. This approach is well motivated because one can view the Plebanski action as a kind of perturbation of the BF action (albeit the perturbation parameter is a Lagrange multiplier field which one needs to integrate over in a path integral). The path integral for BF theory, however, is under good control [11] so that one may hope to obtain a valid path integral formulation for GR by functional current derivation methods [12] familiar from ordinary quantum field theory (QFT).

   As we will try to explain in the next section (see also [3]), the quantum implementation of the so-called simplicity constraints of the Plebanski theory, to the best knowledge of the authors, has not yet been achieved to full satisfaction from first principles in these models. They are called simplicity constraints because they enforce the B field of BF theory to be simple, that is, to originate from a tetrad. Clearly, unless the simplicity constraints are properly implemented, the resulting theory has little to do with quantum gravity. An issue to keep in mind is that the solutions to the classical simplicity constraint consist of five sectors, of which two give rise to $\pm$ times the Palatini action, two give rise to $\pm$ times a topological action and one is a degenerate sector. All these sectors are *a priori* included in a sum over Plebanski histories which may or may not be what one wants[4].

---

[4] For a recent proposal to tackle this issue, see [13].

It is appropriate to also mention further constraints in SFM at this point. The construction of these models relies on a simplicial triangulation τ of the differential four-manifold as well as a dual graph τ^*. A recent analysis has shown [14] that freely specifying geometrical data (areas or fluxes) on the faces of τ tends to lead to inconsistent values of the lengths of the edges of τ unless the so-called Regge constraints are imposed, in addition to the simplicity constraints. These constraints are important to be taken care of if one wants to relate SFM to the established theory of Regge calculus [15] and in order to capture the correct semiclassical limit[5]: in fact, Regge calculus is formulated directly in terms of edge lengths, while in SFM one rather works with electrical fluxes or areas; but a typical simplicial triangulation has far more faces than edges in τ so that assigning a length to an edge from given area values may be ambiguous and/or inconsistent. Note that in our approach, on the other hand, since the path integral is explicitly based on the Holst action, there is no necessity to relate it to the Regge action—which for Plebanski's theory is, of course, a challenge.

In fact, one possibility to make progress on the common issues of the standard formulation of SFM is based on a very simple idea which, to the best knowledge of the authors, occurs for the first time in [16]: namely, simply try to formulate the path integral in terms of the Holst action [17] rather than the Plebanski action. Not only is the Holst action a valid starting point for canonical LQG, but also the simplicity constraints are explicitly solved in that one works entirely with tetrads from the beginning. More precisely, the Holst action uses a specific quadratic expression in the tetrads for the B field of BF theory which also depends on the Immirzi parameter [18]. Hence, the Holst action depends on a specific, non-degenerate linear combination of the four non-degenerate solutions of the simplicity constraints (see the next section for details) and is thus at the same time more general and more restricted because the Holst path integral will not sum over the afore-mentioned five sectors of Plebanski's theory. As already mentioned, it is at present debated how the fact that one actually takes a sum over all histories with a mixture of positive and negative Palatini and topological actions affects the semiclassical properties of the Plebanski path integral.

As observed in [16], since the Holst action is quadratic in the tetrads, one can, in principle, integrate out the tetrad in the resulting Gaussian integral. This has been sketched in [16]; however, the expressions given there are far from rigorous. Here we will give a rigorous expression. Also, we will include the correct measure factor [19] resulting from the second class constraints involved in the Holst action and making sure that the path integral qualifies as a reduced phase space quantization of the theory, as has been stressed in [20]. A similar analysis has been carried out for the Plebanski theory in [21]; however, the resulting measure factor is widely ignored in the SFM literature[6]. The result of the Gaussian integral is an interesting determinant that displays the full nonlinearity of Einstein's theory. When translating

---

[5] Note, however, that there is no reason to require such additional constraints to be implemented in the strict context of canonical LQG, where the holonomy, flux, area, triad or length operators labelled by curves or surfaces have no direct physical meaning (only their occurrence in compound operators assembled from them and which are Dirac observables or constraint operators). Moreover, in general, these curves and surfaces do not even relate to any simplicial structure, so there is no triangulation with respect to which one would be interested in relating the lengths of the edges of its one-skeleton to the areas of its surfaces (in fact in order to establish the relation between LQG and SFM it would seem that one needs to include spin network states on all possible boundary graphs in the SFM analysis—except if one follows the philosophy of [26]). Therefore, Regge-like constraints never occur in LQG.

[6] For a recent review of the relation between SFM and canonical quantization, see [22].

the remaining integral over the connection in the partition function into SFM language, that is, sums over vertex, edge and face representations, one sees that our model differs drastically from all current models.

Of course, we also need to introduce an infrared (IR) and ultraviolet (UV) regulator in the form of a finite cell decomposition. Two observations lead us to depart from the usual SFM approach where one works with simplicial cell complexes. The first one is the result [23] which demonstrates that current semiclassical states used in LQG do not assign good classical behaviour to the volume operator [24] of LQG unless the underlying graph has cubic topology (see also [25, 26]). Since the volume operator plays a pivotal role in LQG as it defines triad operators and hence the dynamics, this is a first motivation to consider cubic triangulations of the four manifold, which we coin 'cubulations' (see, e.g., [27] and references therein). Note that the result of [23] implies that current SFM based on simplicial cell complexes do not admit the semiclassical states [28] as boundary states, which could mean that current models may have to be extended to more general triangulations. The second observation is that the original motivation for considering simplicial cell complexes in current SFM comes from their closeness to BF theory. BF theory is a topological quantum field theory (TQFT) and therefore one would like to keep triangulation independence of the BF SFM amplitude. That this is actually true is a celebrated result in BF theory. In particular, in order to keep triangulation independence it is necessary to integrate the B field over the triangles t of the tringulation and the F field over the faces f bounding the loops in a dual graph [27]. However, GR is not a TQFT and therefore the requirement to have triangulation independence is somewhat unclear. Of course, it is natural if one wants to exploit the properties of BF theory but not if one takes a different route as we tend to do here. Hence, if we drop that requirement, then it is much more natural to refrain from considering the dual graph in addition to the triangulation. Working with cubulations also greatly simplifies the realization of gauge invariance in discrete models. In fact, gauge invariance is related to the closure constraint in SFM which is a subtle issue, as we will see in the next section. If one works just with a triangulation and drops the dual graph, then such issues are easy to take care of. Finally, the use of cubulations also fits nicely with the framework of algebraic quantum gravity [26], which in its minimal version is also formulated in terms of algebraic graphs of cubic topology only.

The architecture of this paper is as follows. In section 2, we give a non-technical review of current SFM. We sketch their derivation from the classical Plebanski action focusing on the points where a first-principles argument is missing. These issues will be the motivation for our different route.

In section 3, we derive the Holst spin foam model using cubulations as a UV regulator as motivated above. As this is an exploratory paper only, we will not worry about convergence issues, which will be properly addressed in subsequent works. More precisely, what we compute are tetrad n-point functions. These should contain sufficient information to compute anything of interest in LQG such as graviton scattering amplitudes as in [29] via Lehmann–Zimmermann–Symanzik-like formulae as in ordinary QFT [30], which allows us to reconstruct the S-matrix from symmetric vacuum n-point functions. Of course, how these n-point functions are related to true observables in a diffeomorphism invariant theory is a subtle issue that will be clarified in a separate paper [36]. Here we only give a summary. The n-point functions can be computed in closed form up to a remaining functional integral over the connections. This can be done for either signature of the spacetime metric. At this point, one could invoke SFM techniques and expand the integral using harmonic analysis on

the gauge group. The resulting intertwiner displays a much more complicated structure than in any current SFM. In particular, pictorially speaking one basic building block is an *octagon diagram*, an analytic expression for which could be called the *96 j symbol* in the case of $G = \mathrm{SO}(4)$. However, the *n*-point functions display a certain Wick-like structure as if they came from a Gaussian integral. What makes the theory interacting and obstructs the tetrads from being a generalized free field[7] is the additional functional integral over the connection. In background-dependent QFT, the moments of a Gaussian measure depend on a background-dependent covariance (usually depending on the Laplacian (in the Euclidean setting) and the mass). Our theory behaves similarly, just that due to background independence the covariance itself is a field that must be integrated over. This is similar in spirit to what happens in three dimensions (3D) [31] when coupling GR to point particles: there, when integrating over the gravitational degrees of freedom one ends up with particles moving on a non-commutative geometry. Here, instead of a non-commutative geometry, we obtain an interacting theory of tetrad fields.

In section 4, we conclude and outline the missing tasks that need to be performed before our model can be taken seriously. An interesting result of our analysis is that in the present formulation, which lacks the simplicity constraints of Plebanski's theory, the irreducible representations of $\mathrm{Spin}(p, 4 - p)$ are not forced to be simple. Simple representations, which basically reduce $\mathrm{Spin}(p, 4 - p)$ to an SU(2) subgroup, can only arise if we impose the time gauge which in the classical theory is used in order to reduce the Holst connection to the Ashtekar–Barbero–Immirzi connection, which is a necessary ingredient of the canonical quantization programme. Gauge fixing conditions, of course, naturally arise in any attempt to make formal path integral expressions better behaved and the situation here is similar.

Two appendices treat some simple technical aspects of this work.

Most parts of this paper do not depend on whether the spacetime signature is Lorentzian or Euclidean.

2. Outline of current spin foam models

In this section, we intend to give a brief summary of the developments in SFM with a focus on the derivation of current models from the Plebanski action and the gaps in that derivation. This serves as motivation for the present paper.

To begin with, it is worth mentioning that the classical solutions to the simplicity constraints actually comprise altogether five sectors, namely two topological sectors $B = \pm e \wedge e$, two Palatini sectors $B = *e \wedge e$ (where $*$ denotes the Hodge map with respect to the internal Minkowski or Euclidean metric) and one degenerate sector. In a path integral, a sum over all these sectors will occur while one would expect that one should only include one of the Palatini sectors or maybe a Holst sector $B = *e \wedge e + \frac{1}{\gamma} e \wedge e$ [17] in order to have a path integral for Einstein's theory. Here γ is the Immirzi parameter of LQG [18].

We now sketch the usual 'derivation' of SFM from the Plebanski action:

The Plebanski action is of the form

$$S = \int \mathrm{Tr}(B \wedge F(A) + \Phi \cdot B \wedge B), \qquad (2.1)$$

---

[7] Roughly, a generalized free scalar field is such that all its *n*-point functions are already determined by its two-point function.

IOP Institute of Physics · DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

where Φ is a scalar Lagrange multiplier field with values in the tensor product of two copies of $so(1, 3)$ or $so(4)$ depending on the signature and F is the curvature of the connection A. In a formal path integral formulation, one integrates $\exp(iS)$ over A, B, Φ. Integrating first over Φ we are left with a partition function of the form

$$Z = \int [dA] [dB] \, \delta(C(B)) \, \exp\left(i \int \, \text{Tr}(B \wedge F)\right), \tag{2.2}$$

where $C(B)$ denotes the collection of the simplicity constraints on B. If one were to solve the delta distribution by integration over B, one would obtain the afore-mentioned sum over the five sectors and integral over the tetrad fields. However, this would result in a complicated expression which does not exploit the relation of Plebanski's formulation to BF theory. Thus, rather than doing that, one notes that roughly speaking

$$B \exp\left(i \int \, \text{Tr}(B \wedge F)\right) = \frac{1}{i} \frac{\delta}{\delta F} \exp\left(i \int \, \text{Tr}(B \wedge F)\right). \tag{2.3}$$

Denoting the functional derivative by X one can now formally pull the δ distribution out of the B integral and perform the integration over B, resulting in

$$Z = \int [dA] \, \delta(C(X)) \cdot \delta(F). \tag{2.4}$$

Without the 'operator' $\delta(C(X))$ this would be the formal partition function of BF theory. Thus, one has achieved the goal to preserve the closeness of the theory to BF theory. One should now expand $\delta(F)$ in terms of eigenfunctions of the collection of operators $C(X)$ and keep only the zero eigenfunctions multiplied by $\delta(0)$.

In order to give meaning to those formal expressions, one has to introduce a UV and IR cutoff as is customary in constructive QFT. That is, one considers finite simplicial triangulations τ of the (possibly compact) differential four manifold and dual graphs τ^*. The two forms B are now approximated by integrals $B(t)$ of B over triangles t of τ, while the curvatures F are approximated by holonomies $A(\partial f)$ around the loops ∂f of the faces f dual to the triangles t. One writes $f(t)$ for the face dual to t. The BF action is then discretized by

$$\sum_{t \in \tau} \text{Tr}(A(\partial f(t)) \, B(t)). \tag{2.5}$$

The reason to work with both τ and τ^* is that, in fact,

$$\int_M \text{Tr}(B \wedge F) = \sum_{t \in \tau} \text{Tr}(F(f(t))B(t)) \tag{2.6}$$

is *an exact identity* [27] where $F(f)$ denotes the integral of F over f. This is very convenient, in particular for pure BF theory. The only approximation thus consists in replacing $F(f)$ by $A(\partial f) - 1_G$.

Likewise, the functional derivatives X must be approximated by ordinary derivatives X_t with respect to the variables $F(t) := A(\partial f(t)) - A(\partial f(t))^T$. Note that when defined like that, the X_t are mutually commuting[8]. However, this is not what is done in current models. Rather one replaces X_t by Y_t; the right invariant vector field on the copy of G associated with the variable $A(\partial f(t))$. Upon spin foam quantization, the discrete B variables thus become explicitly

---

[8] See [32] for the exploration of the model with this definition of X_t.

IOP Institute of Physics   ⬧ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

non-commutative[9]. The reason for doing this replacement is that the Y_t have a simpler action on the delta distribution

$$\delta(F) := \prod_{t \in \tau} \delta_G(A(\partial f(t)). \tag{2.7}$$

It is usually justified by saying that $\delta(F)$ has support on $A(\partial f) = 1_G$ and that Y_t, X_t differ by multiplication with holonomies which should be supported at 1_G. However, this argument is certainly not rigorous because the support of the δ distribution can drastically change when acting with differential operators. Moreover, as already mentioned, this substitution comes with a price: while the simplicity constraints in terms of X_t are mutually commuting, those in terms of Y_t are not. In addition, one does not impose all the simplicity constraints but only a subset of them: there are three types: constraints involving (i) the same triangle, (ii) two triangles sharing an edge and (iii) two triangles sharing a vertex. The latter constraint is implied by the so-called closure constraint on tetrahedra T

$$\sum_{t \in T} Y_t = 0 \tag{2.8}$$

(but not vice versa). This constraint looks as if it would be automatically satisfied because it looks like a gauge invariance condition. However, the product of δ distributions (2.7) in $\Delta(F)$ is not annihilated by the closure constraint (2.8)! This is obvious from the fact that the product of δ distributions involves products of the form

$$\prod_{t \in T} \chi_{\pi_t}(A(\partial f(t)), \tag{2.9}$$

where π denotes an irreducible representation of G and χ_π its character. However, there is no gauge invariant intertwiner among the loops $\partial f(t), t \in T$. One usually argues that the closure constraint is taken into account because *after* integrating over A one is only left with gauge invariant intertwiners, but strictly this is wrong *before* integrating[10]. In fact, since integration with the Haar measure always projects out the gauge invariant part, anything can be made gauge invariant this way. We feel that neglecting the third kind of simplicity constraint (implied by taking the closure constraint for granted) makes the model too local. The effect of truly taking the closure constraint into account is also explored in [32].

As already mentioned, even the simplicity constraints of the first two types are anomalous as they imply vanishing volume [3, 10] and fix the above-mentioned intertwiner to be unique (the model has not enough degrees of freedom). This and other investigations have ruled out[11] the Barrett–Crane model [11], which, however, was an important step in the research on SFM because it triggered the development of model-independent mathematical tools. Recent activities in SFM therefore focused on trying to implement the simplicity constraints of the first two types differently. Thus, for instance, the work [5] uses the Master Constraint type of techniques which were developed in a different context [38] in order to treat second class constraints via a Gupta–Bleuer quantization procedure. In [6], one refrains from imposing the simplicity constraints as operator conditions altogether but rather imposes them semiclassically

---

[9] In fact, as shown in recent work [8, 33, 34], SFM defined as constrained BF models take the form of non-commutative discrete path integrals making use of a star product on functionals of the B variables. It can also be shown that the generating group field theories are just a particular class of non-commutative field theories [35].

[10] See, however, [8, 33, 37] for discussions of this point.

[11] See, however, [33] for a recent critical review of the various arguments raised against the Barrett–Crane model.

by expanding spin foam amplitudes in terms of group coherent states [7] developed by Perelomov [39] and then uses the simplicity condition on the classical bivectors on which the semiclassical amplitudes depend. More recently, the work [9] exploits a spinorial representation of spin network states to implement a Gupta–Bleuer quantization of the simplicity constraints, shown to be solved exactly by coherent states with appropriate labels. Finally, Baratin and Oriti [8, 35] exploit a non-commutative metric representation of spin network states and a non-commutative simplicial path integral representation of quantum BF theory to implement the simplicity constraints as strong constraints on the discrete (and non-commutative) *B* variables.

Some of these methods give rise to models with better semiclassical properties [40] and to better ways to disentangle the topological from the Palatini sector. However, in our opinion a satisfactory derivation from first principles is still missing. By this we mean that one should be able to arrive at those models starting from the Plebanski action, another classically equivalent action or the Hamiltonian formulation and then carry out integrations and imposition of constraints without intermediate approximations or *ad hoc* substitutions as those listed above[12]. This is precisely the motivation of the present paper.

3. Derivation of the model

This is the main section of the paper. It is subdivided into five parts. In the first, we motivate the use of *cubulations* from different perspectives and discuss some of their properties. In the second, we sketch the relation between path integral *n*-point functions and physical (observable) correlators in terms of the physical inner product of the theory. More details on that issue are given in [36]. This crucially works via a choice of gauge fixing or clock system. In the third part, we apply our machinery to *n*-point tetrad functions or equivalently to a generating function of a (complex, regulated) measure. This measure displays a Gaussian-like structure and we can accordingly integrate out half of the degrees of freedom under some assumptions about the choice of gauge fixing. In the fourth part, we discuss the properties of the resulting integral over the remaining degrees of freedom, its Wick-like structure and the structure of the vertex amplitude of the corresponding spin foam model obtained upon harmonic analysis on the gauge group. Finally, in the fifth part we discuss how these *n*-point functions are related to the physical inner product and the kinematical Hilbert space of LQG, in particular how the covariant connection of the Holst path integral reduces to the Ashtekar–Barbero–Immirzi connection of the canonical theory in physical amplitudes.

3.1. Cubulations

In contrast to the standard way to discretize the theory using simplicial triangulations, our approach will be based on *cubulations* of the underlying manifold. The advantages of these, spelled out below, are:

- to facilitate gauge invariant discretization of the classical theory;
- to ensure the existence of a semi-classical sector within the boundary Hilbert space [23].

---

[12] There has been recent work [8, 32, 41] in this direction, where simplicity constraints are clearly implemented in the measure of a path integral. The novelty of the present approach, however, is to start directly from the Holst gravity action, which avoids having to deal with simplicity constraints to begin with.

But the main advantage is a practical one:

- To permit a discretization of the action in terms of a Gaussian with block diagonal kernel, which allows explicit computation of the Gaussian integrals in the partition function.

3.1.1. Gauge invariance. Let us look more closely at the issue of gauge invariance for BF theory which also makes use of a dual graph. Here gauge invariance is not preserved locally (i.e. triangle-wise) in the formula $\int \text{Tr}(B \wedge F) = \sum_t \text{Tr}(B(t)F(f(t)))$ if both B and F transform locally in the adjoint representation. In order to make the gauge transformations more local, one could discretize them. To see how this can be achieved, recall that by the definition of a cell dual to a simplex[13] in a simplicial complex τ, the face $f(t)$ is a union of triangles $[\hat{t}, \hat{T}, \hat{\sigma}]$ subject to the condition $t \subset \partial T, T \subset \partial \sigma$. Here $\hat{(.)}$ denotes the barycentre [27] of a simplex and T, σ denote tetrahedra and four simplices in τ, respectively. So we see that both t and $f(t)$ contain the barycentre $\hat{t}$ in their intersection and we could *define* a disjoint action of the gauge group on both $B(t), F(f(t))$ at $\hat{t}$. However, this is no longer possible when using the approximation $\sum_t \text{Tr}(B(t)A(\partial f(t)))$ because now the only natural action of the gauge group on the loop holonomy is by adjoint action at a starting point on $\partial f(t)$. Now $\partial f(t)$ is a composition of the half-edges $[\hat{T}, \hat{\sigma}]$ where $t \subset \partial T, T \subset \partial \sigma$ but the fundamental degrees of freedom are the holonomies along the edges $e = [\hat{\sigma}, \hat{\sigma}']$ for $\sigma \cap \sigma' = T, t \subset \partial T$. Obviously, the only natural starting point of the loops is then at the vertices $\hat{\sigma}$ which are disjoint from the triangles t. But the triangles are also disjoint from the half-edges as a simple calculation reveals. To maintain gauge invariance, one has to come up with a more complicated discretized action (e.g in terms of wedge variables related to each other by additional holonomy variables [40]). Such complications come from the fact that one is dealing simultaneously with a (simplicial) complex and its dual cell complex; we take this as a further piece of motivation to work only with the triangulation.

3.1.2. Cubulations versus simplicial triangulations. The previous considerations do not specify the type of triangulations to be considered. As already mentioned, the first piece of information why to use cubulations rather than simplicial triangulations is because the boundary graphs must contain cubical ones in order to make sure that the corresponding boundary Hilbert space contains enough semiclassical states [23]. However, there is an additional, more practical motivation for doing so which we discuss now.

Recall that the Holst action is given by

$$S = -\frac{1}{\kappa} \int_M \text{Tr}(G[A] \wedge e \wedge e) = \frac{1}{\kappa} \int_M G_{IJ}[A] \wedge e^I \wedge e^J. \tag{3.1}$$

Here κ denotes Newton's constant,

$$G[A] = 2\left(*F[A] + \frac{1}{\gamma}F[A]\right), \tag{3.2}$$

where $F_{IJ} = dA_{IJ} + A_{IK} \wedge A_J^K$ denotes the curvature of the connection A, γ is the Immirzi parameter, $*$ denotes the internal Hodge dual, that is,

$$(*T)_{IJ} := \frac{1}{2}\epsilon_{IJKL}\eta^{KM}\eta^{LN}T_{MN}, \tag{3.3}$$

---

[13] As usual [27], an n-simplex is denoted by $[p_0, \ldots, p_n]$ where the points p_i denote its corners.

where $I, J, K, \ldots = 0, \ldots, 3$ and η is the Minkowski or Euclidian metric for the structure group $G = \mathrm{SO}(1, 3)$ or $G = \mathrm{SO}(4)$, respectively. As motivated in the introduction, we plan to keep the co-tetrad one forms e^I rather than introducing a B field and thus the simplicity constraints are manifestly solved. Moreover, the issue raised in [14] is circumvented as co-tetrads are labelled by curves and not by (overcomplete) surfaces.

In order to give meaning to a path integral formulation, we consider a UV cutoff in terms of a triangulation τ of M which we choose to be finite, thereby introducing an IR regulator as well. Let us denote the two-dimensional faces of τ by f and the one-dimensional edges of τ by l. We want to discretize (3.1) in a manifestly (and locally) gauge invariant way, just using edges and faces. To do so, we equip all edges with an orientation once and for all. Given an edge l, consider

$$e_l^I := \int_l [A(l(x))]^I {}_J e^J(x). \tag{3.4}$$

Here $l(x)$ for $x \in l$ denotes the segment of l that starts at the starting point of l and ends at x and $[A(p)]^I J$ denotes the G valued holonomy of A along a path p. Evidently, under local gauge transformations $g : M \to G$, (3.4) transforms as $e_l^I \mapsto g_J^I(b(l))e_l^J$ where $b(l)$ denotes the beginning point of l.

To avoid confusion, here $g \in G$ means the following: the fundamental objects are the matrices g_J^I. Set $\tilde{g}_{IJ} := \eta_{IK} g^K {}_J$. Then $g \in G$ iff $\tilde{g}_{IK} \tilde{g}_{JL} \eta^{KL} = \eta_{IJ}$. This is equivalent to $(g^{-1})^I {}_J = \eta^{IL} g^K {}_L \eta_{KJ}$. In other words,

$$\widetilde{(g^{-1})} = (\tilde{g})^{\mathrm{T}}. \tag{3.5}$$

If $g^I {}_J = [\exp(F)]^I {}_J$ for some generator $F^I {}_J$, then (3.5) means that $\tilde{F}_{IJ} + \tilde{F}_{JI} = 0$. In an abuse of notation, one usually uses the same symbols g, F and $\tilde{g}$, $\tilde{F}$, respectively, but unless we are in the Euclidian regime we should pay attention to the index position.

Clearly, the curvature F must be discretized in terms of the holonomy of A along the closed loops ∂f where we have also equipped the faces f with an orientation once and for all. We have

$$
\begin{aligned}
F_{IJ}(f) &:= \frac{1}{2}([\widetilde{A(\partial f)}]_{IJ} - [\widetilde{(A(\partial f))^{-1}}]_{IJ}) \\
&= \frac{1}{2}([\widetilde{A(\partial f)}]_{IJ} - [(\widetilde{(A(\partial f))})^T]_{IJ}) \\
&= \widetilde{A(\partial f)}_{[IJ]} \\
&\approx \int_f F_{IJ}(x),
\end{aligned}
\tag{3.6}
$$

where we have used the non-Abelian Stokes theorem for 'small' loops, that is,

$$A(\partial f) \approx \exp\left(\int_f F\right) \tag{3.7}$$

and we have written $\tilde{F}_{IJ}(x) := F_{IJ}(x)$. We may now define the antisymmetric matrix

$$G_{IJ}(f) = (*F(f))_{IJ} + \frac{1}{\gamma} F_{IJ}(f). \tag{3.8}$$

Imagine now that we were to use a simplicial triangulation. Hence M is a disjoint (up to common tetrahedra) union of four simplices $\sigma = [p_0(\sigma), \ldots, p_4(\sigma)]$. For each $p_j(\sigma)$ label

the four boundary edges of σ starting at $p_j(\sigma)$ by $l_\mu^j(\sigma)$ and let the face (triangle) of σ spanned by $l_\mu^j(\sigma), l_\nu^j(\sigma)$ be denoted by $f_{\mu\nu}^j(\sigma)$ with the convention $f_{\mu\nu}^j(\sigma) = -f_{\nu\mu}^j(\sigma)$. Now the orientation of $l_\mu^j(\sigma)$ either coincides with the given orientation of the corresponding edge in σ or it does not. In the former case define $e_\mu^{Ij}(\sigma) := e_{l_\mu^j(\sigma)}^I$ while in the latter we define $e_\mu^{Ij}(\sigma) := [A(l_\mu^j(\sigma))^{-1} e_{l_\mu^j(\sigma)}]^I$. Then we have

$$\kappa S = -\sum_{\sigma \in \tau} \int_\sigma \mathrm{Tr}(G \wedge e \wedge e)$$

$$\approx \frac{1}{5} \sum_{\sigma \in \tau} \sum_{j=0}^4 \epsilon^{\mu\nu\rho\lambda} G_{IJ}(f_{\mu\nu}^j(\sigma)) e_\rho^{Ij}(\sigma)\, e_\lambda^{Jj}(\sigma)$$

$$=: \sum_{l,l'} G_{IJ}{}^{l,l'} e_l^I\, e_{l'}^J, \tag{3.9}$$

where we have averaged over the corners of a four-simplex. For any simplicial triangulation, the (symmetric in the compound index (I, l)) matrix $G_{IJ}^{ll'}$ is difficult to present explicitly due to bookkeeping problems even if we refrain from averaging over the five corners of a four-simplex. Moreover, as we intend to perform a Gaussian integral over the e_l^I, we need the determinant of that matrix which is impossible to compute explicitly unless it is block diagonal in some sense.

The latter observation points to a possible solution. First of all, any manifold admits a cubulation, that is, a triangulation by embedded hypercubes[14] [42]. We assume that M has a countable cover by open sets O_α. Consider a stratification by four-dimensional (4D) regions S_α subordinate to it. Then S_α admits a regular cubulation, that is, the one-skeleton of the cubulation of M restricted to S_α can be chosen to be a regular cubic lattice. Non-trivial departures from the regular cubulation only appear at the boundaries of the S_α. We restrict our attention to those M admitting a cubulation such that in every compact submanifold the ratio of the number of cubes involved in the non-regular regions divided by the number of cubes involved in the regular regions converges to zero as we take the cubulation to the continuum. For those M, up to corrections which vanish in the continuum limit we can treat M as if it would admit a global, regular cubulation.

Given a regular cubulation τ, consider its set of vertices. In 4D, each vertex v is eight-valent and there are four pairs of edges such that the members of each pair are analytic continuations of each other while the tangents at v of four members from mutually different pairs are linearly independent of each other. It is therefore possible to assign to each edge a direction $\mu = 0, 1, 2, 3$ and an orientation such that adjacent edges in the same direction have a common analytic continuation and agree in their orientation. We label the edges starting at v in the μ direction by $l_\mu(v)$. Note that this labelling exhausts all possible edges and unambiguously assigns an orientation to all of them. The discretized co-tetrad is then given by

$$e_\mu^I(v) := e_{l_\mu(v)}^I. \tag{3.10}$$

---

[14] An easy proof uses the fact that every manifold can be triangulated by simplices. Given a D-simplex, consider the barycentre of each of its $\binom{D+1}{p+1}$ sub-p-simplices for $p = 0, \ldots, D$. Connect the barycentre of any $p+1$-simplex with the barycentres of the p-simplices in its boundary. It is not difficult to see that this defines a cubulation of the D-simplex and that all p-cubes thus defined are the same ones in common q-simplices of the original simplicial complex. In other words, every simplicial complex has a cubulated refinement.

Note that the hypercubic lattice that results solves all our bookkeeping problems since we may now label each vertex by a point in $\mathbb{Z}^4$.

Next, given a vertex v we denote by $v \pm \hat{\mu}$ the next-neighbour vertex in the μ direction. We define the plaquette loop in the μ, ν plane at v by

$$\partial f_{\mu\nu}(v) := l_\mu(v) \circ l_\nu(v + \hat{\mu}) \circ l_\mu(v + \hat{\nu})^{-1} \circ l_\nu(v)^{-1}, \tag{3.11}$$

so that $\partial f_{\nu\mu}(v) = [\partial f_{\mu\nu}(v)]^{-1}$. Note that again this labelling exhausts all minimal loops in one skeleton of τ. The discretized 'curvature' is therefore

$$G_{IJ}^{\mu\nu}(v) := \epsilon^{\mu\nu\rho\sigma} G_{IJ}(f_{\rho\sigma}(v)). \tag{3.12}$$

Denoting by σ the 4D hypercubes in τ we note that there is a one-to-one correspondence between the vertices v in the 0-skeleton of τ and the hypercubes given by assigning to σ that corner $v = (z_0, \ldots, z_3)$ of σ with the smallest values of all $z_0, \ldots, z_3 \in \mathbb{Z}$. We then find that

$$\kappa S = \sum_\sigma \int_M G_{IJ} \wedge e^I \wedge e^J$$

$$\approx \sum_v \sum_{I,J,\mu,\nu} G_{IJ}^{\mu\nu}(v) \, e_\mu^I(v) \, e_\nu^J(v). \tag{3.13}$$

The crucial observation is now the following: assemble pairs of indices into a joint index $A = (I, \mu)$, $B = (J, \nu)$, etc and let $e^A(v) := e_\mu^I(v)$, $G_{AB}(v) := G_{IJ}^{\mu\nu}(v)$, etc. Note that by construction $G_{AB}(v) = G_{BA}(v)$ for all v. Then (3.13) can be written as

$$\kappa S \approx \sum_v e^T(v) \, G(v) \, e(v). \tag{3.14}$$

This means that using (regular) cubulations, indeed the matrix $G_{IJ}^{ll'}$ becomes block diagonal where each block is labelled by a vertex and corresponds to the symmetric 16×16 matrix $G(v)$. This is what makes the computation of the determinant of the huge matrix with entries $G_{IJ}^{ll'}$ practically possible. As we will see, the matrices $G(v)$ have a number of intriguing symmetries, which makes the computation of their determinant an interesting task.

Interesting questions that arise in algebraic topology and which we intend to address in future publications are:

1. Given any D-cubulation, does there exist a cubulated refinement such that one can consistently assign to every D cube σ a vertex v and to all edges an orientation such that there are precisely D edges outgoing from v? We call the cubulations for which this is possible *regular*. If that were the case, we could generalize our discretization from regular hypercubic lattices to arbitrary cubic ones and thus would not have to make any error at the boundaries of the stratified regions mentioned above.

2. If the answer to 1. is negative, can one choose maximally regular cubulations so as to minimize the error in our assumption of globally regular cubulations? In 3D some results on that issue seem to exist [42].

3. Given maximally regular cubulations, can one make an error estimate resulting from the neglect of the non-trivial topology?

3.2. Notes on n-point functions

In the spin foam literature the first task that one addresses is the computation of the partition function. However, the partition function itself has no obvious physical meaning even if one imposes boundary conditions on the paths (spin foams) to be integrated (summed) over. The hope is that SFM provide a formula for the physical inner product of the underlying constrained canonical theory which starts from some kinematical Hilbert space $\mathcal{H}$. The purpose of this section is to sketch the connection between path integrals and n-point functions for a general constrained theory. We will use reduced phase space quantization as our starting point. The connection with operator constraint quantization and group averaging [43] and more details can be found in [36].

We assume that we are given a classical theory with first class constraints $\{F\}$ and possibly second class constraints $\{S\}$. We turn the system into a purely second class system by supplementing $\{F\}$ with suitable gauge fixing conditions $\{G\}$. The canonical Hamiltonian H_c is a linear combination of the primary constraints plus a piece H_0' non-vanishing on the constraint surface of the primary constraints (it could be identically zero). It can also be written as a first class piece H_0 and (some of) the first class constraints F. The gauge fixing conditions fix the Lagrange multipliers involved in the canonical Hamiltonian. One may split the complete set of canonical pairs (q, p) on the full phase space into two sets $(\phi, \pi), (Q, P)$ such that one can solve the system $S = F = G = 0$ which defines the constraint surface for $(\phi, \pi) = f(Q, P)$ in terms of Q, P. The Q, P are coordinates on the reduced phase space which is equipped with the pull-back symplectic structure[15] induced by the embedding of the constraint surface specified by f. The gauge fixing conditions also induce a reduced Hamiltonian H_r which only depends on Q, P and which arises by computing the equations of motion for Q, P with respect to H_c and then restricting them to the gauge fixed values of the Lagrange multipliers and to the constraint surface. Then H_r is defined as the function of Q, P only[16] which generates these same equations of motion. We are now in the situation of an ordinary Hamiltonian system equipped with a true Hamiltonian H_r. We quantize a suitable subalgebra of the reduced Poisson algebra as a $*$ algebra $\mathfrak{A}$ and represent it on a Hilbert space $\mathcal{H}$. This Hilbert space is to be identified with the physical Hilbert space arising from reduced phase space quantization. Its relation with Dirac's constraint quantization is spelled out in [36]. Let $t \mapsto U(t)$ be the unitary evolution induced by H_r. Then the object of interest is the transition amplitude or n-point function

$$\langle \psi_f, U(t_f - t_n)a_n U(t_n - t_{n-1})a_{n-1}\ldots U(t_2 - t_1)a_1 U(t_1 - t_i)\psi_i \rangle \qquad (3.15)$$

between initial and final states ψ_i, ψ_f at initial and final times t_i, t_f, respectively, with intermediate measurements of the operators $a_1, \ldots, a_n \in \mathfrak{A}$ at $t_1 < t_2 < \cdots < t_n$.

Preferably one would like to be in a situation in which there is a cyclic vector Ω for $\mathfrak{A}$, which is also a ground state for H_r. The existence of a cyclic vector is no restriction because representations of $\mathfrak{A}$ are always direct sums of cyclic representations. In this case $\mathfrak{A}\Omega$ is dense in $\mathcal{H}$ and we may therefore restrict attention to $\psi_i = \psi_f = \Omega$ by choosing appropriate $a_1, \ldots, a_n$ in (3.15). The existence of a vacuum state for H_r means that zero is in the point spectrum of H_r. Let us make this assumption for simplicity.

Let us abbreviate the Heisenberg time evolution as $a_k(t) := U(t)^{-1}a_k U(t)$. In principle, it would be sufficient to restrict the a_k to be configuration operators Q because their time evolution

---

[15] This symplectic structure coincides with the pull-back of the degenerate symplectic structure on the full phase space corresponding to the Dirac bracket induced by the system $\{S, F, G\}$ [44].

[16] For simplicity, we are assuming a gauge fixing which leads to a conservative reduced Hamiltonian.

contains sufficient information about P as well. However, we will stick to the more general case for reasons that will become clear later. This leads us the n-point function

$$S(t_1, \ldots, t_n) := \frac{\langle \Omega, U(t_f) a_n(t_n) \ldots a_1(t_1) U(-t_i) \Omega \rangle}{\langle \Omega, U(t_f - t_i) \Omega \rangle}, \qquad (3.16)$$

where we have properly normalized so as to give the zero-point function the value unity. This has the advantage that certain infinities that would otherwise arise in the following can be absorbed. Note that since Ω is a ground state, the $U(t_f)$ and $U(t_i)$ as well as the denominator could be dropped in (3.16).

Now a combination of well-known heuristic arguments [44, 45] reviewed in [36] reveals the following: consider any initial and final configuration q_i, q_f on the full phase space and denote by $\mathcal{P}((t_i, q_i), (t_f, q_f))$ the set of paths[17] in full configuration space between (q_i, q_f) at times t_i, t_f, respectively. Consider

$$Z[j; q_i, q_f] = \lim_{-t_i, t_f \to \infty} \int_{\mathcal{P}((t_i, q_i), (t_f, q_f))} [Dq \, Dp \, D\lambda \, D\mu] \delta[G] |\det[\{F, G\}]| \rho \, e^{\frac{i}{\hbar} S[q, p, \lambda, \mu]} \cdot e^{i \int_{t_i}^{t_f} dt \, j(t) \cdot q(t)}$$

$$(3.17)$$

Here j is a current in the fibre bundle dual to that of q, $S[q, p, \lambda, \mu]$ is the canonical action after performing the singular Legendre transform from the Lagrangian to the Hamiltonian formulation[18] and ρ is a local function of q, p which is usually related to the Dirac bracket determinant $\det[\{S, S\}]$ [45].

Now the primary constraints are always of the form $\pi = f(Q, P, \phi)$ where we have split again the canonical pairs into two groups. Thus, $S[q, p, \lambda, \mu]$ is linear in those momenta π and we can integrate them out yielding δ distributions of the form $\delta[\lambda - (.)]\delta[\mu - (.)]$, which can be solved by integrating over λ, μ. If we assume that the dependence of the remaining action on P is only quadratic and that G and $|\det[\{F, G\}]|$ are independent of P then we can integrate also over P which yields in general a Jacobian I coming from the Legendre transform. We can then write (3.17) as

$$Z[j; q_i, q_f] = \lim_{-t_i, t_f \to \infty} \int_{\mathcal{P}((t_i, c_i), (t_f, c_f))} [Dq] \, \delta[G] \, |\det[\{F, G\}]| \, \rho \, I \, e^{\frac{i}{\hbar} S[q]} e^{i \int_{t_i}^{t_f} dt \, j(t) \cdot q(t)}, \qquad (3.18)$$

where proper substitutions of π from solving the primary constraints and of P from the Legendre transformation are understood. Here $S[q]$ is the original (covariant) Lagrangian action.

Defining $\chi[j] := \frac{Z[j; q_i, q_f]}{Z[0; q_i, q_f]}$ the covariant or path integral n-point functions

$$S(t_1, \ldots, t_n) := \left[\frac{\delta^n \chi[j]}{i^n \delta j(t_1) \ldots \delta j(t_n)} \right]_{j=0} \qquad (3.19)$$

have the canonical or physical interpretation of

$$\langle \Omega, T(a_1(t_1) \ldots a_n(t_n)) \Omega \rangle, \qquad (3.20)$$

where T is the time ordering symbol, Ω is the afore-mentioned cyclic vacuum vector defined by the physical (or reduced) Hamiltonian H_r induced by the gauge fixing G, $a_k(t)$ is the Heisenberg

---

[17] This should be a suitable measurable space but we leave it unspecified.

[18] The Lagrange multipliers λ, μ of the primary first and second class constraints, respectively, play the role of the velocities that one could not solve for in terms of the momenta in the process of the Legendre transform.

operator at time t (evolved with respect to H_r) corresponding to a_k, and a_k itself classically corresponds to a component of q evaluated on the constraint surface $S = F = G = 0$. The scalar product corresponds to a quantization on the reduced phase space defined by G. Note how the gauge fixing condition G (or choice of clocks) prominently finds its way both into the canonical theory and into the path integral formula (3.18). In particular, note that the seemingly similar expression

$$Z'[j; q_i, q_f] = \lim_{-t_i, t_f \to \infty} \int_{\mathcal{P}((t_i,q_i),(t_f,q_f))} [Dq] \, e^{\frac{i}{\hbar} S[q]} \, e^{i \int_{t_i}^{t_f} dt \, j(t) \cdot q(t)} \tag{3.21}$$

does not have any obvious physical interpretation and in addition lacks the important measure factors ρ, I.

Remarks:

1. One may be puzzled by the following: from ordinary gauge theories on background spacetimes such as the Yang–Mills theory on Minkowski space, the path integral or more precisely the generating functional of the Schwinger functions (in the Euclidian formulation) does not require any gauge fixing in order to give the path integral a physical interpretation. One needs it only in order to divide out the gauge volume in a systematic way (Fadeev–Popov identity) while the generating functional is independent of the gauge fixing. The gauge fixing also does not enter the construction of gauge invariant functions (such as Wilson loops). In our case, however, the gauge fixing condition is actually needed in order to formulate the physical time evolution and the preferred choice of gauge invariant functions on phase space.

   The resolution is as follows: the difference between the Yang–Mills theory and generally covariant systems such as GR that we are interested in here is indeed that the canonical Hamiltonian is, in fact, the generator of gauge transformations (spacetime diffeomorphisms) rather than physical time evolution. It is even constrained to vanish. In contrast, in the Yang–Mills theory there is a preferred and gauge invariant Hamiltonian which is not constrained to vanish. In order to equip the theory with a notion of time, we have used the relational framework discovered in [46] which consists in choosing fields as clocks and rods with respect to which other fields evolve. Mathematically, this is equivalent to a choice of gauge fixing. Hence, in our case the gauge fixing plays a dual role: firstly, to render the generating functional less singular, and secondly, to define physical time evolution.

2. The appearance of the δ distributions and functional (Fadeev–Popov) determinants in (3.17) indicates that we are not dealing with an ordinary Hamiltonian system but rather with a constrained system. One can, in fact, get rid of the gauge fixing condition involved if one pays a price. The price is that if one considers instead of q its gauge invariant extension $\tilde{q}$ off the surface $G = 0$ [44, 47], then, since we consider the quotient $Z[j]/Z[0]$ which leads to *connected n-point functions*, by the usual Fadeev–Popov identity that exploits gauge invariance, we may replace [36] (3.17) by

$$\tilde{Z}[j, q_i, q_f] = \int_{\mathcal{P}((t_i,q_i),(t_f,q_f))} [Dq \, Dp \, D\lambda \, D\mu] \, \rho \, e^{\frac{i}{\hbar} S[q,p,\lambda,\mu]} \, e^{i \int_{t_i}^{t_f} dt \, j(t) \cdot \tilde{q}(t)}. \tag{3.22}$$

   However, (3.22) is not very useful unless $\tilde{q}(q, p)$ is easy to calculate which is typically not the case. Hence, we will refrain from doing so. Nevertheless, no matter whether one deals with (3.17) or (3.22), the correlation functions depend on the gauge fixing G or in other

words on the choice of the clocks [47, 48] with respect to which one defines a physical reference system.

3. The correspondence between (3.19) and (3.20) also allows us to reconstruct the physical inner product from the n-point functions: given arbitrary states $\psi, \psi' \in \mathcal{H}$ we find $a, a' \in \mathfrak{A}$ such that $||a\Omega - \psi||, ||a'\Omega - \psi'||$ are arbitrarily small. Now pick any $t_i < t_0 < t_f$, then

$$\langle a\Omega, a'\Omega \rangle = \langle \Omega, a^\dagger a'\Omega \rangle, \tag{3.23}$$

By assumption, the operator $a^\dagger a'$ can be written as a finite linear combination of monomials of homogeneous degree in the components of the operator q which we write, suppressing indices for the components as q^n. Then

$$\langle \Omega, q^n\Omega \rangle = \lim_{t_1,\ldots,t_n \to t_0; t_n > \cdots > t_1} \langle \Omega, q(t_n)\ldots q(t_1)\Omega \rangle, \tag{3.24}$$

which can be expressed via (3.19). The existence of this coincidence limit of n-point functions is often problematic in background-dependent Wightman QFT [30], but their existence is actually the starting point of canonical quantization of background-independent non-Wightman QFT as one can see from the identity (3.24).

3.3. The generating functional of tetrad n-point functions

We now want to apply the general framework of the previous section to GR in the Holst formulation. Classically, it is clear that without fermions all the geometry is encoded in the co-tetrad fields e_μ^I because then the spacetime connection is just the spin connection defined by the co-tetrad (on shell). If fermions are coupled, the same is still true in the second-order formulation so that there is no torsion. But even in the first-order formulation with torsion one can attribute the torsion to the fermionic degrees of freedom. Hence we want to consider as a complete list of configuration fields the co-tetrad.

We will now make two assumptions about the choice of gauge fixing and the matter content of our system:

I. The local measure factors ρ, I depend on the co-tetrad only analytically. This is actually true for the Holst action [19], see also [21].

II. The gauge fixing condition G is independent of the co-tetrad and the Fadeev–Popov determinant $\det(\{F, G\})$ depends only analytically on the co-tetrad. With respect to the first class Hamiltonian and spatial diffeomorphism constraint, this can always be achieved by choosing suitable matter as a reference system, see, e.g., [49]. However, in addition there is the Gauss-law first class constraint. Here it is customary to impose the time gauge condition [17] which asks that certain components of the tetrad vanish. This will also enable one to make the connection with canonical LQG where one works in the time gauge in order to arrive at an SU(2) rather than G connection. Fortunately, in this case it is possible to explicitly construct a complete set of G-invariant functions of the tetrad, namely the four metric[19] $g_{\mu\nu} = e_\mu^I e_\nu^J \eta_{IJ}$, and if we only consider correlators of those, then we can get rid of the time gauge condition as indicated in the previous section (Fadeev–Popov identity). In section 3.5, we will come back to this issue, however, in trying to make the connection of the SFM obtained with canonical LQG for which the time gauge is unavoidable. We will then sketch how to possibly relax the assumptions made under II.

[19] In the presence of fermions there are additional gauge invariant functions also involving the fermions.

Under the assumptions made we consider the generating functional $\chi(j, J) := Z[j, J]/Z[0, 0]$ where

$$
Z[j, J] := \int [D\phi \; DA] \, \delta[G[A, \phi]] \, \mathrm{e}^{\mathrm{i} \int_M \mathrm{Tr}(J \wedge \phi)}
$$
$$
\times \int [De] \, \rho[e, A, \phi] \, I[e, A, \phi] \, |\det[\{F, G\}]|[e, A, \phi] \, \mathrm{e}^{\frac{\mathrm{i}}{\hbar}(S_g[e,A]+S_m[e,A,\phi])} \mathrm{e}^{\mathrm{i} \int_M \mathrm{Tr}(j \wedge e)}.
$$
(3.25)

Here ϕ denotes the matter configuration variable. We have split the total action into the geometry (Holst) part S_{g} and a matter part S_{m} which typically depends non-trivially but analytically on e. Also the total current was split into pieces J, j, respectively, taking values in the bundles dual to those of ϕ, e, respectively.

A confusing and peculiar feature of first-order actions such as the Holst or Palatini action is that, from a Lagrangian point of view, both fields e, A must be considered as configuration variables. In performing the Legendre transform [19], one discovers that there are primary constraints which relate certain combinations of e to the momenta conjugate to A. One can solve these constraints and then (A, e) appear as momentum and configuration coordinates of this partly reduced phase space. This is the reason why we consider only correlations with respect to e. The idea is now as is usual in path integral theory: we set

$$
\sigma[e, A, \phi] := \rho[e, A, \phi] \, I[e, A, \phi] \, |\det[\{F, G\}]|[e, A, \phi] \, \mathrm{e}^{\frac{\mathrm{i}}{\hbar} S_{\mathrm{m}}[e,A,\phi]}
$$
(3.26)

and write (3.25) as

$$
Z[j, J] := \int [D\phi \, DA] \, \delta[G[A, \phi]] \, \mathrm{e}^{\mathrm{i} \int_M \mathrm{Tr}(J \wedge \phi)} \, \sigma\left[\frac{\delta}{\mathrm{i}\delta j}, A, \phi\right] \left\{ \int [De] \, \mathrm{e}^{\frac{\mathrm{i}}{\hbar} S_{\mathrm{g}}[e,A]} \mathrm{e}^{\mathrm{i} \int_M \mathrm{Tr}(j \wedge e)} \right\}.
$$
(3.27)

Of course, $\mathrm{e}^{\mathrm{i} S_{\mathrm{m}}}$ must be power-expanded in a perturbation series in order to carry out the functional derivations with respect to j. Indeed, if we consider just the functional integration with respect to e and think of A, ϕ as external fields, then the piece S_{g} being quadratic in e is like a *free part*, while S_{m} being only analytic in e is like an *interaction part* of the action as far as the co-tetrad is concerned. Of course, in the computation of the physical tetrad n-point functions all the functional derivatives involved in (3.27) are eventually evaluated at $j = 0$.

It follows that the object of ultimate interest is the *Gaussian integral*

$$
z[j; A] := \int [De] \, \mathrm{e}^{-\frac{\mathrm{i}}{\ell_P^2} \int_M \mathrm{Tr}(G \wedge e \wedge e + \ell_P^2 \, j \wedge e)},
$$
(3.28)

which is computable exactly. Of course, it is not a standard Gaussian, firstly because the exponent is purely imaginary and secondly because the 'metric' $G_{IJ}^{\mu\nu}(A)$ is indefinite so that $z[j; A]$ would be ill-defined if the exponent was real[20]. In the appendix, we recall how to integrate such non-standard Gaussians. In order to carry out this integral we must make the

---

[20] As is usual this prevents a 'Euclidian' version of GR. Here Euclidian stands for Euclidian field theory with an analytic continuation of the imaginary axis of the real time variable involved (Wick rotation), which leads to a real exponent. This has nothing to do with Lorentzian or Euclidian signature GR. In fact, most metrics do not have an analytic section so that Wick rotation is ill-defined and thus the connection between the real and the Euclidian theory is unclear.

technical assumption that configurations A for which G is singular have measure zero with respect to DA. We will come back to this assumption later.

It is at this point where we must regularize the path integral in order to perform the Gaussian integration[21] and we write the discretized version on a cubulation of M as motivated in section 3.1; that is, we replace (3.28) by the discretized version

$$z[j; A] := \int \prod_{v,I,\mu} \mathrm{d}e_\mu^I(v) \, \mathrm{e}^{\frac{\mathrm{i}}{\ell_P^2} \sum_v [G_{IJ}^{\mu\nu}(v)e_\mu^I(v)e_\nu^J(v) + \ell_P^2 j_\mu^I(v)e_\mu^I(v)]}. \tag{3.29}$$

The results of appendices A and B now reveal that

$$z[j; A] := \left[\prod_v \frac{\mathrm{e}^{\frac{\mathrm{i}\pi}{4}\mathrm{ind}(G(v))}}{\sqrt{|\det(G(v))|}} \right] \mathrm{e}^{-\mathrm{i}\frac{\ell_P^2}{4} \sum_v [G^{-1}(v)]_{\mu\nu}^{IJ} j_I^\mu(v) j_J^\nu(v)}, \tag{3.30}$$

where we dropped a factor $\sqrt{\pi}^{16N}$ for a cubulation with N vertices because it is cancelled by the same factor coming from the denominator in $\chi(j, J)$, see (3.25).

3.4. Wick structure, graviton propagator and spin foam model vertex structure

3.4.1. Wick structure. Formula (3.30) explicitly displays the main result of our investigation: the full j dependence of the generating functional written as (3.27) rests in (3.30). We are interested in the nth functional derivatives of (3.30) at $j = 0$. Now similarly to free field theories, the corresponding n-point functions vanish for n odd. However, in contrast to free field theories, for n even, the n-point functions cannot be written in terms of polynomials of the two-point function. The reason is that the 'covariance' $G^{-1}[A]$ of the Gaussian is not a background structure but rather depends on the quantum field A itself that one has to integrate over. This renders the co-tetrad theory a non-quasi-free, that is, interacting theory. Nevertheless, it is true that all Wick identities that have been derived for free field theories still hold also for the n-point tetrad functions *albeit in the sense of expectation values or means with respect to A.*

3.4.2. Graviton propagator. To illustrate this, consider a fictive theory in which $\sigma(e, A, \phi)$, $G(A, \phi)$ are both independent of both A, e. This is not a very physical assumption but serves to make some observations of general validity in a simplified context. This means that we can drop the ϕ dependence because the generating functional factorizes. Thus in our fictive theory, we are looking at the generating functional $\chi[j] = z[j]/z[0]$, where

$$z[j] = \int [DA] \, z[j; A] = \int \prod_{v,\mu} \mathrm{d}\mu_\mathrm{H}(A(l_\mu(v))) \left[\prod_v \frac{\mathrm{e}^{\frac{\mathrm{i}\pi}{4}\mathrm{ind}(G(v))}}{\sqrt{|\det(G(v))|}} \right] \mathrm{e}^{-\frac{\mathrm{i}\ell_P^2}{4} \sum_v j_I^\mu(v) \, j_J^\nu(v), [G^{-1}(v)]_{\mu\nu}^{IJ}} \tag{3.31}$$

where μ_H is the[22] Haar measure on G. Now let

$$\langle e_{\mu_1}^{I_1}(v_1) \dots e_{\mu_n}^{I_n}(v_n) \rangle := \left[\frac{\delta^n \chi[j]}{\mathrm{i}^n \delta j_{I_1}^{\mu_1}(v_1) \cdots \delta j_{I_1}^{\mu_1}(v_1)} \right]_{j=0}. \tag{3.32}$$

---

[21] Actually, we can formally solve the Gaussian integral w/o specifying the triangulation, i.e. we can compute it in the continuum. However, then one must regularize the resulting determinant which amounts to the same problem.
[22] In the case of non-compact G the Haar measure is unique up to a normalization constant which drops out in $\chi(j)$. To choose the Haar measure instead of the Lebesgue measure makes sense in the continuum limit of infinitely 'short' edges as usual.

IOP Institute of Physics **Φ** DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

It is immediately clear that

$$\langle e^{I_1}_{\mu_1}(v_1)\, e^{I_2}_{\mu_2}(v_2)\rangle = 0 \tag{3.33}$$

unless $v_1 = v_2$. This is reassuring because as we said above, physically it only makes sense to consider correlators of G-invariant objects such as the metric. The simplest n-point function of interest is therefore the four-point function

$$\langle g_{\mu_1\nu_1}(v_1)\, g_{\mu_2\nu_2}(v_2)\rangle = \langle e^{I_1}_{\mu_1}(v_1)\, e^{J_1}_{\nu_1}(v_1) e^{I_2}_{\mu_2}(v_2)\, e^{J_2}_{\nu_2}(v_2)\rangle\, \eta_{I_1 J_2}\, \eta_{I_2 J_2}. \tag{3.34}$$

If we are interested in something like a graviton propagator, we are interested in $v_1 \neq v_2$ and obtain

$$\langle g_{\mu_1\nu_1}(v_1)\, g_{\mu_2\nu_2}(v_2)\rangle = \left[\frac{\ell_P^2}{2}\right]^4 \langle [G(v_1)^{-1}]^{I_1 J_1}_{\mu_1\nu_1}\, [G(v_2)^{-1}]^{I_2 J_2}_{\mu_2\nu_2}\rangle', \tag{3.35}$$

where for $F = F[A]$

$$\langle (F)\rangle' := \frac{\int \prod_{v,\mu}\, \mathrm{d}\mu_{\mathrm{H}}(A(l_\mu(v))\left[\prod_v \frac{\mathrm{e}^{\frac{\mathrm{i}\pi}{4}\,\mathrm{ind}(G(v))}}{\sqrt{|\det(G(v))|}}\right] F[A]}{\int \prod_{v,\mu}\, \mathrm{d}\mu_{\mathrm{H}}(A(l_\mu(v))\left[\prod_v \frac{\mathrm{e}^{\frac{\mathrm{i}\pi}{4}\,\mathrm{ind}(G(v))}}{\sqrt{|\det(G(v))|}}\right]}. \tag{3.36}$$

Note that $G(v)^{-1}$ does not share the symmetries of $G(v)$, so $[G^{-1}(v)]^{(IJ)}_{\mu\nu}$ does not vanish automatically. We see that we are basically interested in correlators of the inverse matrix $G(v)^{-1}$ with respect to the joint Haar measure. Whether these have the correct behaviour in a situation where, instead of vacuum boundary states, one chooses coherent states peaked on a classical background metric as suggested in [25, 29] is currently under investigation.

3.4.3. Spin foam vertex structure. Finally, in order to translate (3.36) into spin foam language, we should perform harmonic analysis on G and write the integrand of the Haar measure in terms of irreducible representations of G. In particular, the vertex structure of a spin foam is encoded in $z[0]$ so that we are interested in harmonic analysis of the function

$$F(v) := \frac{\mathrm{e}^{\frac{\mathrm{i}\pi}{4}\,\mathrm{ind}(G(v))}}{\sqrt{\det(G(v))|}}. \tag{3.37}$$

To derive its graph theoretical structure it is enough to find which $F(v)$ depend on a given holonomy $A(l)$. Recall that $F(v)$ is a function cylindrical over the graph $\gamma(v) = \cup_{\mu<\nu}\partial f_{\mu\nu}(v)$, which is the union of its respective plaquette loops. Consider a fixed edge $l = l_\mu(v)$. It is contained in $\gamma(v')$ if and only if it is contained in one of the plaquette loops $\partial f_{\mu\nu}(v')$ or $\partial f_{\nu\mu}(v')$ with $\mu < \nu$ or $\nu < \mu$, respectively. In the first case it must coincide either with $l_\mu(v')$ or with $l_\mu(v'+\hat{v})$. In the second case it must coincide either with $l_\mu(v'+\hat{v})$ or with $l_\mu(v')$ as well. Thus in either case we must have either $v' = v$ or $v' = v - \hat{v}$, $v \neq \mu$.

For our illustration purposes, let us consider for simplicity that G is compact, the non-compact case has the same spin foam vertex structure but the harmonic analysis is slightly more complicated. Then each function $F(v)$ can be formally expanded into SO(4) (or rather the universal cover SU(2) × SU(2)) irreducible representations[23] with respect to the six plaquette holonomies $A(\partial f_{\mu\nu}(v))$, $\mu < \nu$. These representations π are labelled by pairs of half-integral

---

[23] This expansion would be rigorous if we knew that $F(v)$ is an L_2 function that is currently under investigation. We assume here that in any case we may use the Peter and Weyl theorem in a distributional sense.

spin quantum numbers but we will not need this for what follows. Thus $F(v)$ admits an expansion of the form

$$F(v) = \sum_{\{\pi_{\mu\nu}\}} \iota'_{\{\pi_{\mu\nu}\}} \cdot [\otimes_{\mu<\nu} \pi_{\mu\nu}(A(\partial f_{\mu\nu}(v)))], \tag{3.38}$$

where $\iota'_{\{\pi_{\mu\nu}\}}$ is a gauge invariant intertwiner for the sextuple of irreducible representations $\{\pi_{\mu\nu}\}_{\mu<\nu}$ which is independent of v, the only v dependence rests in the holonomies. It depends on the specific algebraic form of $F(v)$ which derives from the Holst action.

Let us define $\pi_{\nu\mu} := \pi_{\mu\nu}$ for $\mu < \nu$. By writing the six plaquette holonomies in terms of four edge holonomies, it is not difficult to see that $F(v)$ can also be written in the form

$$F(v) = \sum_{\{\pi_{\mu\nu}\}} \iota_{\{\pi_{\mu\nu}\}} \left[\underset{\mu,\mu\neq\nu}{\otimes} \pi_{\mu\nu}(A(l_\mu(v))) \otimes \pi_{\mu\nu} A(l_\mu(v+\hat{v})) \right], \tag{3.39}$$

which displays explicitly the 16 variables $A(l_\mu(v))$, $A(l_\mu(v+\hat{v})$, $\nu \neq \mu$ involved and consists of $24 = 6 \times 4$ tensor product factors. In order to arrive at (3.39) we had to rearrange the contraction indices which induces the change from ι' to ι and we have made use of $\pi(A(l)^{-1}) = \pi^{\mathrm{T}}(A(l))$ for $G = \mathrm{SO}(4)$.

We may now carry out explicitly the integrals over edge holonomies in $z[0]$ by inserting the expansion (3.39). We write symbolically[24]

$$z[0] = \int \prod_{v,\mu} \mathrm{d}\mu_{\mathrm{H}}(A(l_\mu(v))) \prod_{v'} F(v')$$

$$= \sum_{\{\pi^v_{\mu\nu}\}} \left[\prod_v \iota_{\{\pi^v_{\mu\nu}\}} \cdot \right] \int \prod_{v,\mu} \mathrm{d}\mu_{\mathrm{H}}(A(l_\mu(v))) [\underset{v',\mu,\mu\neq\nu}{\otimes} \pi^{v'}_{\mu\nu}(A(l_\mu(v'))) \otimes \pi^{v'}_{\mu\nu}(A(l_\mu(v'+\hat{v}))]$$

$$= \sum_{\{\pi^v_{\mu\nu}\}} \left[\prod_v \iota_{\{\pi^v_{\mu\nu}\}} \cdot \right] \int \prod_{v,\mu} \mathrm{d}\mu_{\mathrm{H}}(A(l_\mu(v))) [\underset{v',\mu,\mu\neq\nu}{\otimes} \pi^{v'}_{\mu\nu}(A(l_\mu(v'))) \otimes \pi^{v'-\hat{v}}_{\mu\nu}(A(l_\mu(v'))]$$

$$= \sum_{\{\pi^v_{\mu\nu}\}} \left[\prod_v \iota_{\{\pi^v_{\mu\nu}\}} \cdot \right] \underset{v,\mu}{\otimes} \left[\int_G \mathrm{d}\mu_{\mathrm{H}}(g) [\underset{\mu\neq\nu}{\otimes} \pi^v_{\mu\nu}(g) \otimes \pi^{v-\hat{v}}_{\mu\nu}(g) \right]. \tag{3.40}$$

Here in the second step we have shifted the vertex label in one of the tensor product factors in order to bring out the dependence on the $A(l_\mu(v))$. It follows that the end result of the integration is that for each edge $l = l_\mu(v)$ there is a gauge invariant intertwiner[25]

$$\rho_{\{\pi^v_{\mu\nu}, \pi^{v-\hat{v}}_{\mu\nu}\}_{\nu\neq\mu}} := \left[\int_G \mathrm{d}\mu_{\mathrm{H}}(g) [\underset{\mu\neq\nu}{\otimes} \pi^v_{\mu\nu}(g) \otimes \pi^{v-\hat{v}}_{\mu\nu}(g) \right], \tag{3.41}$$

which intertwines *six* representations rather than four as in (constrained) BF theory on simplicial triangulations. The origin of this discrepancy is, of course, that we are using cubulations rather than simplicial triangulations. These six representations involved for edge $l_\mu(v)$ correspond

---

[24] We rearrange the tensor products as if they were scalars but this can be corrected by performing corresponding rearrangements in the contraction structure of the intertwiners. We assume this to be done without explicitly keeping track of it because it does not change the vertex structure.

[25] We do not bother to expand it into a recoupling scheme and thus to label the intertwiner itself by three irreducible representations.
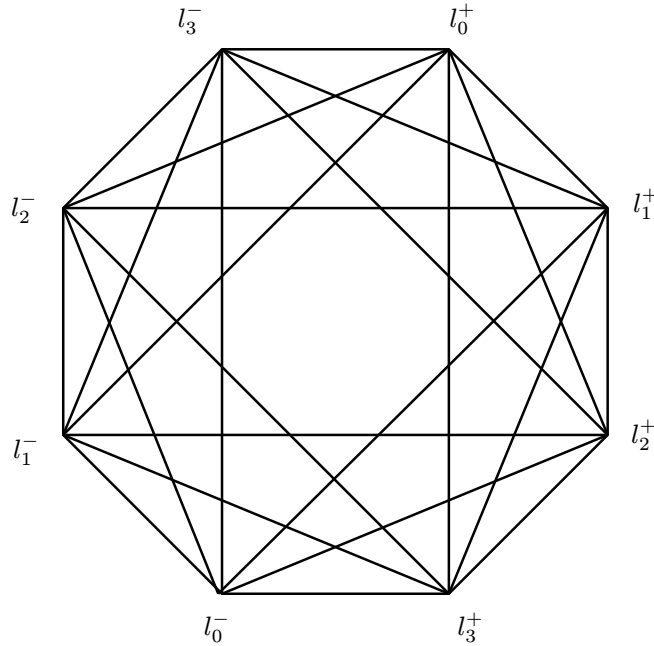
Figure 1. The octagon diagram associated with vertex v. The eight corners correspond to the eight edges $l = l_\mu^\sigma(v) = l_\mu(v + \frac{\sigma-1}{2}\hat\mu)$, $\sigma = \pm$ adjacent to v. The line between corners labelled by $l_\mu^\sigma(v), l_\nu^{\sigma'}(v)$ for $\mu \neq \nu$ corresponds to the face $f = f_{\mu\nu}^{\sigma\sigma'}(v) = f_{\mu\nu}(v + \frac{\sigma-1}{2}\hat\mu + \frac{\sigma'-1}{2}\hat\nu)$. We should colour corners by intertwiners ρ_l and lines by representations π_f but refrain from doing so in order not to clutter the diagram. Altogether 48 irreducible representations of Spin(4) (or 96 of SU(2)) are involved.

precisely to the six plaquette loops $\partial f_{\mu\nu}(v)$, $\partial f_{\mu\nu}(v - \hat\nu)$, $\nu \neq \mu$ of which $l_\mu(v)$ is a segment. Therefore, if we associate with each face $f = f_{\mu\nu}(v)$ an irreducible representation $\pi_f = \pi_{\mu\nu}^v$ and denote by $\{\pi\}$ the collection of all the π_f, then the basic building block (3.41) can be written in the more compact form

$$\rho_l[\{\pi\}] = \int_G \mathrm{d}\mu_\mathrm{H}(g) \bigotimes_{l \subset \partial f} \pi_f(g). \tag{3.42}$$

Likewise, if we denote $\iota_v[\{\pi\}] := \iota_{\{\pi_{\mu\nu}^v\}}$, then

$$z[0] = \sum_{\{\pi\}} \left[\prod_v \iota_v[\{\pi\}]\cdot \right] \bigotimes_l \rho_l[\{\pi\}], \tag{3.43}$$

which, of course, hides the precise tensor product and contraction structure but is sufficient for our purposes.

Formula (3.43) is precisely the general structure of an SFM. Moreover, the intertwiner (3.42) is the direct analogue of the intertwiner in BF theory which there defines the pentagon diagram [11]. If we were to try and draw a corresponding picture for our model, then for each vertex v we would draw eight points, one for each edge l incident at v. These edges are labelled by the intertwiner ρ_l. Given two points corresponding to edges l, l' consider the unique face f that has l, l' in its boundary. Draw a line between each such points and label it by π_f. The result is the *octagon diagram*, see figure 1. Concretely, the edges adjacent to v are

$l_\mu(v)$, $l_\mu(v - \hat\mu)$, $\mu = 0, 1, 2, 3$. For $\mu \neq \nu$, the face spanned by $l_\mu(v), l_\nu(v)$ is $f_{\mu\nu}(v)$, the face spanned by $l_\mu(v)$, $l_\nu(v - \hat\nu)$ is $f_{\mu\nu}(v - \hat\nu)$, the face spanned by $l_\mu(v - \hat\mu), l_\nu(v)$ is $f_{\mu\nu}(v - \hat\mu)$ and finally the face spanned by $l_\mu(v - \hat\mu), l_\nu(v - \hat\nu)$ is $f_{\mu\nu}(v - \hat\mu - \hat\nu)$. The corresponding labels on the lines are thus $\pi_{\mu\nu}^v$, $\pi_{\mu\nu}^{v-\hat\nu}$, $\pi_{\mu\nu}^{v-\hat\mu}$, $\pi_{\mu\nu}^{v-\hat\mu-\hat\nu}$, respectively. Thus the octagon diagram has eight points and $6 \times 4 = 24$ lines (each line connects two points). These correspond to the 24 plaquettes that have a corner in v; namely for each $\mu < \nu$ these are $f_{\mu\nu}(v)$, $f_{\mu\nu}(v - \hat\mu)$, $f_{\mu\nu}(v - \hat\nu)$, $f_{\mu\nu}(v - \hat\mu - \hat\mu)$. In the case of $G = \mathrm{SO}(4)$ each irreducible representation is labelled by two spin quantum numbers. The intertwiner freedom is labelled by three irreducible representations of $\mathrm{SO}(4)$ and there is one irreducible representation corresponding to a face. Thus the octagon diagram depends on $3 \times 8 + 24 = 48$ irreps of $\mathrm{SO}(4)$ or 96 spin quantum numbers. Since each intertwiner (3.42) factorizes into *two* intertwiners [3] (one for the starting point and one for the beginning point of the edge but both depending on the same representations) we may actually collect those eight intertwiners associated with the same vertex. The collection of those eight factors is actually the analytic expression corresponding to the octagon diagram which therefore may be called the *96 j symbol*.

The decisive difference between (constrained) BF theory and our model is however that in (constrained) BF theory the analogue of the function $F(v)$ is a product of δ distributions, one for each face holonomy. The simplicity constraints just impose restrictions on the representations and intertwiners, but this cannot change the fact that there is factorization in the face dependence. In our model, the face dependence does not factorize; hence, in this sense it is less local or more interacting.

To sum up this section: we have explicitly described the analytical expression for the vertex amplitude of this SFM in (3.40)–(3.43), which leads to the octagon diagram described. Using harmonic analysis on $\mathrm{SO}(4) = \mathrm{SU}(2) \times \mathrm{SU}(2)/\mathbb{Z}_2$, one can easily describe everything in terms of spin representations of $\mathrm{SU}(2)$. As the resulting expression is not very illuminating, we refrain from displaying it here.

3.5. Relation between covariant and canonical connection

Another striking feature of our model is the following: constrained BF theory, that is, the Plebanski theory, should be a candidate for quantum gravity. Our Holst model should be equivalent to that theory at least semiclassically because morally speaking, the only difference between them lies in the technical implementation of the simplicity constraints, modulo the caveats mentioned in section 2. Now one of the most important properties of the implementation of the simplicity constraints in some of the most popular SFM is that the irreducible Spin(4) representations that one sums over are the simple ones[26]. In our model we do not see any sign of that[27]. This is an important issue because the restriction to simple representations means that the underlying gauge theory is roughly SU(2) rather than Spin(4), which looks correct if the SFM is to arise from canonical LQG, which indeed is a SU(2) gauge theory. Thus, in usual SFM the simplicity constraints seem to already imply the gauge fixing of the 'boost' part of the Spin(4) Gauss constraint that is necessary at the classical level in order to pass from the Holst connection to the Ashtekar–Barbero–Immirzi connection [17]. Strictly speaking, that has

---

[26] If we label an irrep of Spin(4) by a pair (j_+, j_-), then a simple irrep is one for which $j_+ = j_-$ [11]. There is a similar restriction if one works with an arbitrary Immirzi parameter [5].

[27] Also the models proposed in [8, 9, 32] do not restrict to simple representations. The amplitudes are however peaked on these, although with a non-trivial width.

not yet been established as pointed out in [50], where it is shown that the connection used in SFM is actually the spin connection and not the Holst connection. But apart from that, in the considerations of the previous section we do not see any simplicity restrictions on the type of group representations.

However, note that what we did in the previous section was incomplete because in order to properly define the n-point functions we must gauge fix the generating functional with respect to the G Gauss constraints. Formally, this is not necessary if we only consider correlators of G invariant functions such as the metric because the infinite gauge group volume formally cancels out in the fraction $z[j]/z[0]$. However, details matter: the formal arguments cannot be substantiated by hard proofs in this case. Specifically, if we consider $G = \mathrm{SO}(1, 3)$, there is no measure known for gauge theories for non-compact groups (see [51] for the occurring complications) and thus we are forced to gauge fix at least the boost part of the Gauss constraint. This is the same reason why one uses the time gauge in the canonical theory. We expect that implementing the time gauge fixing [17] in a way similar to the implementation of the simplicity constraints in usual BF theory will effectively reduce the gauge group to $\mathrm{SU}(2)$.

Details will appear elsewhere, but roughly speaking the idea is the following: the time gauge is a set of constraints $C[e]$ on the co-tetrad e. By the usual manipulations we can pull the corresponding δ distribution out of the co-tetrad fuctional integral and formally obtain

$$\chi_{\mathrm{ABI}}[j] = [\delta[C[\delta/\delta j]]\, \chi_{\mathrm{Holst}}[j], \tag{3.44}$$

where χ_{Holst} is the generating functional of the previous section and χ_{ABI} stands for the Ashtekar–Barbero–Immirzi path integral.

Whether this really works in a rigorous fashion remains to be seen. However, we find it puzzling that the simplicity constraints in usual SFM, which classically have nothing to do with the time gauge, should automatically yield the correct boundary Hilbert space. It seems intuitively clear that the time gauge must be imposed in the quantum theory in addition to the simplicity constraints, just like in the classical theory, as we suggest. Without imposing it, we do not see any sign of a restriction from G to $\mathrm{SU}(2)$ in our model where we solve the simplicity constraints differently. We expect this to be related to the work [52]. This observation indicates that usual SFM and our formulation are rather different from each other.

4. Conclusions and future work

In this paper we have proposed a different strategy to construct SFM for GR. Rather than the Plebanski action we take the Holst action as our starting point. This means that the simplicity constraints of the Plebanski formulation have been correctly taken care of. The price to pay is that the connection to BF theory is lost. The motivation behind our strategy is that BF theory is a TQFT and therefore quite different from GR which has an infinite number of physical degrees of freedom. Hence the usual triangulation-independent methods developed for TQFT and employed in current SFM are possibly less powerful in the context of GR. In particular, the fact that it is difficult to deal with the simplicity constraints in current SFM might be a sign of that. Another problem with the Plebanski formulation that we have not mentioned yet is that it is difficult to couple matter because matter directly couples to the co-tetrad rather than the B field. In principle, one can express e via B modulo simplicity constraints but the corresponding formulae are even more involved than those for e. Note that one *must* couple matter to BF theory in order to get a realistic model. For 3D gravity, coupling matter is straightforward [53] because

IOP Institute of Physics Φ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

there the *B* field and *e* coincide, while in 4D this has not yet been done except for non-standard model fermions which just couple to the connection [54, 55] or membranes coupled to pure BF theory [56].

The method we propose in this paper might be called a brute force and textbook strategy. Dropping the insistence on triangulation independence right from the beginning, we proposed a Wilson action-like naive discretization of the Holst action. We carefully studied the connection of the Holst path integral with the canonical LQG correlation or *n*-point functions and used relational techniques to make the connection with the physical Hilbert space and observables. In principle, none of these ingredients are new; they have been successfully employed in other contexts. Of course, the appealing elegance of (constrained) BF theory has disappeared in our formulation, the integrals to be computed are rather challenging (but not impossible) and the gauge fixing conditions for spatial diffeomorphism and Hamiltonian constraints as well as a local measure factor without which the connection to observation is lost complicate the formalism. However, we feel that the resulting structure, while far from being worked out in detail, has some interesting features such as the Wick-like structure of the physical tetrad correlators and less local structure. In particular, we have shown that imposing the time gauge also in the quantum theory comes out as a necessary and natural condition in our model in order to make contact with the LQG Hilbert space.

As already mentioned in the introduction, this paper is exploratory in nature. It focuses more on ideas rather than analysis and there are many open issues that need to be settled before the present model can be taken seriously. Apart from the topological issues mentioned in section 3.1, the convergence and measure theoretic issues discussed in section 3.3 and finally the issues with the imposition of the time gauge outlined in section 3.5, there are further points that need to be addressed.

One of the most serious ones is the continuum limit: the fact that we are working with cubulations suggests a naive but natural notion of continuum limit, which consists in studying the behaviour of the correlation functions under barycentric refinement of the hypercubes at a fixed IR regulator (boundary surface). The last couple of years has seen the resurgence of a coarse graining programme in SFM (see, e.g., [57]); and it would be very interesting to see how the application of the procedure to our discretization compares with the standard models. Of course, in the spirit of the AQG framework [26], one could also say that the continuum limit has already been taken provided that one works with infinite cubulations. This requires then, in a separate step, to remove the IR regulator.

A more practical but still important problem is the following: even at a finite UV and IR regulator, it is already hard enough to compute the determinant of the covariance matrix of the co-tetrad Gaussian and to determine its index (which may vanish automatically, see appendix B). But since these covariances are highly correlated, practical computation of the *n*-point functions at least in the macroscopic regime will be possible only if the corresponding non-trivial measure has some kind of cluster property [58].

All these issues are left for future work.

Acknowledgments

Appendix A. Non-standard Gaussian integrals

Let G be a real valued, symmetric and non-singular matrix on the real vector space $V = \mathbb{R}^n$ and let $j \in V$. We are interested in the non-standard Gaussian integral

$$I := \int_V \mathrm{d}^n x \; \mathrm{e}^{\mathrm{i}[\frac{1}{2} x^\mathrm{T} G x + j^\mathrm{T} x]}. \tag{A.1}$$

Using the translation $x = y - G^{-1} j$, we can simplify this to

$$I = \mathrm{e}^{-\mathrm{i}\frac{i}{2} j^\mathrm{T} G^{-1} j} \int_V \mathrm{d}^n y \; \mathrm{e}^{\frac{i}{2} y^\mathrm{T} G y}. \tag{A.2}$$

There exists an element $S \in GL(n, \mathbb{R})$ such that $G = S^\mathrm{T} D S$, where D is a regular diagonal $n \times n$ matrix which is possibly indefinite. Denote by $d_1, \ldots, d_n \in \mathbb{R} - \{0\}$ the entries of D. Then the change of coordinates $y = Sz$ reveals

$$I = \frac{\mathrm{e}^{-\mathrm{i}\frac{i}{2} j^\mathrm{T} G^{-1} j}}{|\det(S)|} \int_V \mathrm{d}^n z \; \mathrm{e}^{\frac{i}{2} z^\mathrm{T} D z}. \tag{A.3}$$

Now consider the basic integral

$$I_d := \int_{\mathbb{R}} \mathrm{d}z \; \mathrm{e}^{\mathrm{i}\, d z^2 / 2} \tag{A.4}$$

for $d \in \mathbb{R} - \{0\}$. For $d = \mathrm{i}k$, $k > 0$ we know the value of (A.4); however, that formula involves a square root and thus analytic continuation of I_d in d is ambiguous. Hence we must determine the value of (A.4) by an independent means.

The integrand in (A.4) is entirely analytic in z without poles. For $d > 0$ or $d < 0$, respectively, the integral over the arc $0 \leqslant \arg(z) \leqslant \pi/2$ or $0 \geqslant \arg(z) \geqslant -\pi/2$, respectively, vanishes at infinite radius. Hence, using a Cauchy integral argument, we may rotate the integral from $z \in \mathbb{R}$ to $z \in \mathrm{e}^{\mathrm{i}\mathrm{sgn}(d)\pi/4}\mathbb{R}$ so that with $z = \mathrm{e}^{\mathrm{i}\mathrm{sgn}(d)\pi/4}t$, $t \in \mathbb{R}$, we obtain

$$I_d = \mathrm{e}^{\mathrm{i}\mathrm{sgn}(d)\pi/4} \int_{\mathbb{R}} \mathrm{d}t \; \mathrm{e}^{-|d|t^2/2} = \sqrt{2\pi/|d|} \; \mathrm{e}^{\mathrm{i}\mathrm{sgn}(d)\pi/4}. \tag{A.5}$$

Given a symmetric matrix G with signature p, q (i.e. p positive, q negative and $n - p - q$ zero eigenvalues) we define its *index* $\mathrm{ind}(G) := p - q$. Then, combining (A.3) and (A.5) we obtain

$$I = \frac{\sqrt{2\pi}^n \, \mathrm{e}^{-\mathrm{i}\frac{i}{2} j^\mathrm{T} G^{-1} j} \, \mathrm{e}^{\mathrm{i}\,\mathrm{ind}(G)\pi/4}}{|\det(G)|}. \tag{A.6}$$

Appendix B. On the index of special matrices

While determinants may be tedious to calculate, it is always analytically possible. However, the index of a matrix is harder to obtain. While there exist algorithms to obtain it just from its characteristic polynomial (rather than from its spectrum, which would be impossible to determine analytically for general large matrices) for *concrete matrices* such as the Déscartes sign rule [59], for general matrices of a given restricted structure there are no such algorithms available except for a few cases.

Our situation is the following: consider the 16×16 matrix G with entries $G_{(\mu I),(\nu J)} := G_{IJ}^{\mu\nu}$. Since $G_{IJ}^{\mu\nu} = -G_{IJ}^{\nu\mu} = -G_{JI}^{\mu\nu} = G_{JI}^{\nu\mu}$ it is symmetric. Let us also write $e^{\mu I} := e_{\mu}^{I}$. We consider the lexicographic ordering of the compound index (μI) as $(00), \ldots, (03), (10), \ldots, (13), (20), \ldots, (23), (30), \ldots, (33)$. Consider the antisymmetric 4×4 matrix $G^{\mu\nu}$ with $0 \leqslant \mu < \nu \leqslant 3$ given by $(G^{\mu\nu})_{IJ} := G_{IJ}^{\mu\nu}$. Then the 16×16 matrix G has the following block structure:

$$G = \begin{pmatrix} 0 & G^{01} & G^{02} & G^{03} \\ -G^{01} & 0 & G^{12} & G^{13} \\ -G^{02} & -G^{12} & 0 & G^{23} \\ -G^{03} & -G^{13} & -G^{23} & 0 \end{pmatrix}. \tag{B.1}$$

This leads us to the following conjecture.

Conjecture B.1. *Let A, B, C, D, E, F be antisymmetric, real valued 4×4 matrices and let G be the symmetric 16×16 matrix*

$$G = \begin{pmatrix} 0 & A & B & C \\ -A & 0 & D & E \\ -B & -D & 0 & F \\ -C & -E & -F & 0 \end{pmatrix}. \tag{B.2}$$

Then $\mathrm{ind}(G) = 0$.

It turns out to be extremely hard to prove this conjecture although it is rather plausible. For example, it is easy to show that the conjecture is correct when the matrices A, B, C, D, E and F are 2×2 antisymmetric matrices. It is also true when the matrices A, B, C, D, E and F are linearly dependent. We leave the proof (or disproof) of this conjecture to future publications.

If the conjecture were true and G is non singular, then we will know that G has eight positive and eight negative eigenvalues. Hence we will know that $\det(G) > 0$. In order to compute $\det(G)$ we make use of the following basic factorization property for an arbitrary block matrix with blocks A, B, C, D with invertible A

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ C & 1 \end{pmatrix} \begin{pmatrix} 1 & A^{-1}B \\ 0 & D - CA^{-1}B \end{pmatrix}. \tag{B.3}$$

It follows that

$$\det(G) = \det(A) \, \det(D - CA^{-1}B). \tag{B.4}$$

In our situation, by means of (B.4) we can iteratively downsize the size of the matrix of which we have to compute the determinant from rank 16 to 8 and then to 4. At rank 4 we may use Cayley's theorem [59] in order to express $\det(G)$ directly in terms of polynomials of the traces of products of the $G^{\mu\nu}$ and thus in terms of traces of products of the plaquette loops.

References

[1] Perez A 2003 Spin foam models for quantum gravity *Class. Quantum Grav.* **20** R43
   Baez J C 2000 An introduction to spin foam models of quantum gravity and BF theory *Lect. Notes Phys.* **543** 25–94
[2] Rovelli C 2004 *Quantum Gravity* (Cambridge: Cambridge University Press)

[3] Thiemann T 2006 *Modern Canonical Quantum General Relativity* (Cambridge: Cambridge University Press) arXiv:gr-qc/0110034

[4] Ashtekar A and Lewandowski J 2004 Background independent quantum gravity: a status report *Class. Quantum Grav.* **21** R53

Smolin L 2004 An invitation to loop quantum gravity arXiv:hep-th/0408048

Thiemann T 2003 Lectures on loop quantum gravity *Lect. Notes Phys.* **631** 41–135

Rovelli C 2011 Zakopane lectures on loop gravity arXiv:1102.3660 [gr-qc]

[5] Engle J, Pereira R and Rovelli C 2008 Flipped spinfoam vertex and loop gravity *Nucl. Phys.* B **798** 251–90

Engle J, Livine E, Pereira R and Rovelli C 2008 LQG vertex with finite Immirzi parameter *Nucl. Phys.* B **799** 136–49

[6] Freidel L and Krasnov K 2008 A new spin foam model for 4D gravity *Class. Quantum Grav.* **25** 125018

[7] Livine E R and Speziale S 2007 A new spinfoam vertex for quantum gravity *Phys. Rev.* D **76** 084028

Livine E R and Speziale S 2008 Consistently solving the simplicity constraints for spinfoam quantum gravity *Europhys. Lett.* **81** 50004

[8] Baratin A and Oriti D 2012 Group field theory and simplicial gravity path integrals: a model for Holst–Plebanski gravity *Phys. Rev.* D **85** 044003

[9] Livine E and Dupuis M 2011 Holomorphic simplicity constraints for 4D spinfoam models *Class. Quantum Grav.* **28** 215022

[10] Perez A 2002 Spin foam quantisation of $SO(4)$ Plebanski's action *Adv. Theor. Math. Phys.* **5** 947–68

[11] Crane L and Yetter D A 1993 Categorical construction of 4D TQFTs *Quantum Topology* ed L Kauffman and R Baadhio (Singapore: World Scientific) pp 120–30

Crane L, Kauffman L and Yetter D 1997 State-sum invariants of 4-manifolds *J. Knot Theory Ramifications* **6** 177–234

Barrett J W and Crane L 1998 Relativistic spin networks and quantum gravity *J. Math. Phys.* **39** 3296–302

Barrett J W and Crane L 2000 A Lorentzian signature model for quantum general relativity *Class. Quantum Grav.* **17** 3101–18

[12] Freidel L and Krasnov K 1999 Spin foam models and the classical action principle *Adv. Theor. Math. Phys.* **2** 1183–247

[13] Engle J 2012 A spin-foam vertex amplitude with the correct semiclassical limit arXiv:1201.2187 [gr-qc]

[14] Dittrich B and Ryan J P 2011 Phase space descriptions for simplicial 4D geometries *Class. Quantum Grav.* **28** 065006

[15] Williams R M and Tuckey P 1992 Regge calculus: a bibliography and brief review *Class. Quantum Grav.* **9** 1409–22

Williams R M 1997 Recent progress in Regge calculus *Nucl. Phys. Proc. Suppl.* **57** 73–81

[16] Mikovic A 2005 Tetrade spin foam model arXiv:r-qc/0504131

Mikovic A 2006 Spin foam models from the tetrad integration *AIP Conf. Proc.* **841** 523–7

[17] Holst S 1996 Barbero's Hamiltonian derived from a generalised Hilbert–Palatini action *Phys. Rev.* D **53** 5966

Barros E Sá N 2001 Hamiltonian analysis of general relativity with the Immirzi parameter *Int. J. Mod. Phys.* D **10** 261–72

[18] Immirzi G 1997 Quantum gravity and Regge calculus *Nucl. Phys. Proc. Suppl.* **57** 65

[19] Engle J, Han M and Thiemann T 2010 Canonical path integral measures for Holst and Plebanski gravity. I. Reduced phase space derivation *Class. Quantum Grav.* **27** 245014

Han M 2010 Canonical path-integral measures for Holst and Plebanski gravity: II. Gauge invariance and physical inner product *Class. Quantum Grav.* **27** 245015

[20] Bojowald M and Perez A 2010 Spin foam quantisation and anomalies *Gen. Rel. Grav.* **42** 877–907

[21] Buffenoir E, Henneaux M, Noui K and Roche Ph 2004 Hamiltonian analysis of Plebanski theory *Class. Quantum Grav.* **21** 5203–20

[22] Alexandrov S, Geiller M and Noui K 2011 Spin foams and canonical quantization arXiv:1112.196 [gr-qc]

[23] Flori C and Thiemann T 2008 Semiclassical analysis of the loop quantum gravity volume operator: I. Flux coherent states arXiv:0812.1537 [gr-qc]
      Flori C 2009 Semiclassical analysis of the loop quantum gravity volume operator: II. Area coherent states arXiv:0904.1303 [gr-qc]

[24] Rovelli C and Smolin L 1995 Discreteness of volume and area in quantum gravity *Nucl. Phys.* B **456** 753
      Rovelli C and Smolin L 1995 Discreteness of volume and area in quantum gravity *Nucl. Phys.* B **442** 593–622 (erratum)
      Ashtekar A and Lewandowski J 1997 Quantum theory of geometry. I: Area operators *Class. Quantum Grav.* **14** A55–82

[25] Sahlmann H and Thiemann T 2006 Towards the QFT on curved spacetime limit of QGR. 1. A general scheme *Class. Quantum Grav.* **23** 867–908
      Sahlmann H and Thiemann T 2006 Towards the QFT on curved spacetime limit of QGR. 2. A concrete implementation *Class. Quantum Grav.* **23** 909–54
      Brunnemann J and Thiemann T 2006 Simplification of the spectral analysis of the volume operator in loop quantum gravity *Class. Quantum Grav.* **23** 1289–346

[26] Giesel K and Thiemann T 2007 Algebraic quantum gravity (AQG). I. Conceptual setup *Class. Quantum Grav.* **24** 2465–98
      Giesel K and Thiemann T 2007 Algebraic quantum gravity (AQG). II. Semiclassical analysis *Class. Quantum Grav.* **24** 2499–564
      Giesel K and Thiemann T 2007 Algebraic quantum gravity (AQG). III. Semiclassical perturbation theory *Class. Quantum Grav.* **24** 2565–88
      Giesel K and Thiemann T 2007 Algebraic quantum gravity (AQG). IV. Reduced phase space quantisation of loop quantum gravity arXiv:0711.0119 [gr-qc]

[27] Whitney H 1957 *Geometric Integration Theory* (Princeton, NJ: Princeton University Press)

[28] Thiemann T 2006 Complexifier coherent states for canonical quantum general relativity *Class. Quantum Grav.* **23** 2063–118
      Thiemann T 2001 Gauge field theory coherent states (GCS): I. General properties *Class. Quantum Grav.* **18** 2025–64
      Thiemann T and Winkler O 2001 Gauge field theory coherent states (GCS): II. Peakedness properties *Class. Quantum Grav.* **18** 2561–636
      Thiemann T and Winkler O 2001 Gauge field theory coherent states (GCS): III. Ehrenfest theorems *Class. Quantum Grav.* **18** 4629–81
      Thiemann T and Winkler O 2001 Gauge field theory coherent states (GCS): IV. Infinite tensor product and thermodynamic limit *Class. Quantum Grav.* **18** 4997–5033
      Sahlmann H, Thiemann T and Winkler O 2001 Coherent states for canonical quantum general relativity and the infinite tensor product extension *Nucl. Phys.* B **606** 401–40

[29] Rovelli C 2006 Graviton propagator from background-independent quantum gravity *Phys. Rev. Lett.* **97** 151301
      Bianchi E, Modesto L, Rovelli S and Speziale S 2006 Graviton propagator in loop quantum gravity *Class. Quantum Grav.* **23** 6989–7028
      Alesci E and Rovelli C 2007 The complete LQG propagator. I. Difficulties with the Barrett–Crane vertex *Phys. Rev.* D **76** 104012
      Alesci E and Rovelli C 2008 The complete LQG propagator. II. Asymptotic behavior of the vertex *Phys. Rev.* D **77** 044024
      Alesci E, Bianchi E and Rovelli C 2009 LQG propagator: III. The new vertex *Class. Quantum Grav.* **26** 215001

[30] Haag R 1996 *Local Quantum Physics* 2nd edn (Berlin: Springer)

[31] Freidel L and Livine E 2006 Ponzano–Regge model revisited. III: Feynman diagrams and effective field theory *Class. Quantum Grav.* **23** 2021–62

Freidel L and Livine E 2006 3D quantum gravity and effective non-commutative quantum field theory *Phys. Rev. Lett.* **96** 221301

[32] Han M and Thiemann T 2010 Commuting simplicity and closure constraints for 4D spin foam models arXiv:1010.5444 [gr-qc]

[33] Baratin A and Oriti D 2011 Quantum simplicial geometry in the group field theory formalism: reconsidering the Barrett–Crane model *New J. Phys.* **13** 125011

[34] Baratin A, Dittrich B, Oriti D and Tambornino J 2011 Non-commutative flux representation for loop quantum gravity *Class. Quantum Grav.* **28** 175011

[35] Baratin A and Oriti D 2010 Group field theory with non-commutative metric variables *Phys. Rev. Lett.* **105** 221302

[36] Han M and Thiemann T 2010 On the relation between operator constraint-, master constraint-, reduced phase space- and path integral quantization *Class. Quantum Grav.* **27** 225019

[37] Conrady F and Freidel L 2009 Quantum geometry from phase space reduction *J. Math. Phys.* **50** 123510

[38] Thiemann T 2006 The phoenix project: master constraint programme for loop quantum gravity *Class. Quantum Grav.* **23** 2211–48

Thiemann T 2006 Quantum spin dynamics (QSD): VIII. The master constraint *Class. Quantum Grav.* **23** 2249–66

[39] Perelomov A 1986 *Generalised Coherent States and their Applications* (Berlin: Springer)

[40] Conrady F and Freidel L 2008 Path integral representation of spin foam models of 4D gravity *Class. Quantum Grav.* **25** 245010

Conrady F and Freidel L 2008 On the semiclassical limit of 4D spin foam models *Phys. Rev.* D **78** 104023

[41] Bonzom V and Livine E R 2009 A Lagrangian approach to the Barrett–Crane spin foam model *Phys. Rev.* D **79** 064034

Bonzom V 2009 Spin foam models for quantum gravity from lattice path integrals *Phys. Rev.* D **80** 064028

[42] Martins J F and Mikovic A 2008 Spin foam perturbation theory for three-dimensional quantum gravity arXiv:0804.2811 [gr-qc]

[43] Giulini D and Marolf D 1999 On the generality of refined algebraic quantization *Class. Quantum Grav.* **16** 2479–88

Giulini D and Marolf D 1999 A uniqueness theorem for constraint quantization *Class. Quantum Grav.* **16** 2489–505

[44] Henneaux M and Teitelboim C 1992 *Quantisation of Gauge Systems* (Princeton, NJ: Princeton University Press)

[45] Henneaux M and Slavnov A 1994 A note on the path integral for systems with primary and secondary second class constraints *Phys. Lett.* B **338** 47–50

[46] Rovelli C 1991 What is observable in classical and quantum gravity? *Class. Quantum Grav.* **8** 297–316

Rovelli C 1991 Quantum reference systems *Class. Quantum Grav.* **8** 317–32

Rovelli C 1991 Time in quantum gravity: physics beyond the Schrödinger regime *Phys. Rev.* D **43** 442–56

Rovelli C 1990 Quantum mechanics without time: a model *Phys. Rev.* D **42** 2638–46

[47] Dittrich B 2007 Partial and complete observables for Hamiltonian constrained systems *Gen. Rel. Grav.* **39** 1891–927

Dittrich B 2006 Partial and complete observables for canonical general relativity *Class. Quantum Grav.* **23** 6155–84

[48] Thiemann T 2006 Reduced phase space quantization and Dirac observables *Class. Quantum Grav.* **23** 1163–80

[49] Giesel K, Hofmann S, Thiemann T and Winkler O 2010 Manifestly gauge-invariant general relativistic perturbation theory foundations: I *Class. Quantum Grav.* **27** 055005

Giesel K, Hofmann S, Thiemann T and Winkler O 2010 Manifestly gauge-invariant general relativistic perturbation theory: II. FRW background and first order *Class. Quantum Grav.* **27** 055006

IOP Institute of Physics ⊕ DEUTSCHE PHYSIKALISCHE GESELLSCHAFT

[50] Alexandrov S 2008 Simplicity and closure constraints in spin foam models of gravity *Phys. Rev.* D **78** 044033
Alexandrov S 2008 Spin foam model from canonical quantization *Phys. Rev.* D **77** 024009
[51] Freidel L and Livine E R 2003 Spin networks for noncompact groups *J. Math. Phys.* **44** 1322–56
[52] Livine E 2002 Projected spin networks for Lorentz connection: linking spin foams and loop gravity *Class. Quantum Grav.* **19** 5525–42
[53] Freidel L and Louapre D 2004 Ponzano–Regge model revisited. I: Gauge fixing, observables and interacting spinning particles *Class. Quantum Grav.* **21** 5685–726
[54] Mikovic A 2002 Spin foam models of matter coupled to gravity *Class. Quantum Grav.* **19** 2335–54
[55] Bianchi E, Han M, Magliaro E, Perini C, Rovelli C and Wieland W 2010 Spin foam fermions arXiv:1012.4719 [gr-qc]
[56] Baez J and Perez A 2007 Quantization of strings and branes coupled to BF theory *Adv. Theor. Math. Phys.* **11** 3
Fairbairn W and Perez A 2008 Extended matter coupled to BF theory *Phys. Rev.* D **78** 024013
[57] Dittrich B, Eckert F C and Martin-Benito M 2011 Coarse graining methods for spin net and spin foam models arXiv:1109.4927 [gr-qc]
Dittrich B 2012 From the discrete to the continuous—towards a cylindrically consistent dynamics arXiv:1205.6127 [gr-qc]
[58] Glimm J and Jaffe A 1987 *Quantum Physics* (New York: Springer)
[59] Fischer G 1985 *Analytische Geometrie* (Braunschweig: Vieweg)