On the implementation of the canonical quantum simplicity constraint

N Bodendorfer\textsuperscript{1,2}, T Thiemann\textsuperscript{1,3} and A Thurn\textsuperscript{1}

\textsuperscript{1} Institute for Theoretical Physics III, FAU Erlangen—Nürnberg, Staudtstr. 7, Erlangen D-91058, Germany
\textsuperscript{2} Institute for Gravitation and the Cosmos and Physics Department, Penn State, University Park, PA 16802, USA
\textsuperscript{3} Perimeter Institute for Theoretical Physics, 31 Caroline Street N, Waterloo, ON N2 L 2Y5, Canada

E-mail: norbert.bodendorfer@gravity.fau.de, thomas.thiemann@gravity.fau.de, tthiemann@perimeterinstitute.ca and andreas.thurn@gravity.fau.de

Received 24 January 2012, in final form 11 December 2012
Published 23 January 2013
Online at stacks.iop.org/CQG/30/045005

Abstract
In this paper, we discuss several approaches to solve the quadratic and linear simplicity constraints in the context of the canonical formulations of higher dimensional general relativity and supergravity developed in our companion papers. Since the canonical quadratic simplicity constraint operators have been shown to be anomalous in any dimension $D \geq 3$ in \textit{Class. Quantum Grav.} \textbf{30} 045003, non-standard methods have to be employed to avoid inconsistencies in the quantum theory. We show that one can choose a subset of quadratic simplicity constraint operators which are non-anomalous among themselves and allow for a natural unitary map of the spin networks in the kernel of these simplicity constraint operators to the $\text{SU}(2)$-based Ashtekar–Lewandowski Hilbert space in $D = 3$. The linear constraint operators on the other hand are non-anomalous by themselves; however, their solution space is shown to differ in $D = 3$ from the expected Ashtekar–Lewandowski Hilbert space. We comment on possible strategies to make a connection to the quadratic theory. Also, we comment on the relation of our proposals to the existing work in the spin foam literature and how these works could be used in the canonical theory. We emphasize that many ideas developed in this paper are certainly incomplete and should be considered as suggestions for possible starting points for more satisfactory treatments in the future.

PACS numbers: 04.60.Pp, 04.20.Fy, 04.50.−h
1. Introduction

In [1, 2], gravity in any dimension $D + 1 \geq 3$ has been formulated as a gauge theory of SO$(1, D)$ or of the compact group SO$(D + 1)$, irrespective of the spacetime signature. The resulting theory has been obtained by two different routes: a Hamiltonian analysis of the Palatini action making use of the procedure of gauge unfixing\(^4\), and on the canonical side by an extension of the ADM phase space. The additional constraints appearing in this formulation, the simplicity constraints, are well known. They constrain bivectors to be simple, i.e. the antisymmetrized product of two vectors. Originally introduced in Plebanski’s [10] formulation of general relativity as a constrained BF theory in $(3 + 1)$ dimensions, they have been generalized to an arbitrary dimension in [11] and were considered in the context of Hamiltonian lattice gravity [12, 13]. Moreover, discrete versions of the simplicity constraints are a standard ingredient of the spin foam approaches to quantum gravity [14–16], see [17, 18] for reviews, and recently were also used in group field theory [19–21] as well as on a simplicial phase space [22, 23], where also their algebra was calculated. Two different versions of simplicity constraints are considered in the literature, which are either quadratic or linear in the bivector fields. The quantum operators corresponding to the quadratic simplicity constraints have been found to be anomalous both in the covariant [24] as well as in the canonical picture [25, 3]. On the covariant side, this lead to one of the major points of critique about the Barrett–Crane model [14]: the anomalous constraints are imposed strongly\(^5\), which may imply erroneous elimination of physical degrees of freedom [26]. This triggered the development of the new spin foam models [27, 28, 15, 24, 16, 29], in which the quadratic simplicity constraints are replaced by linear simplicity constraints. The linear version of the constraints is slightly stronger than the quadratic constraints, since in $(3 + 1)$ dimensions the topological solution is absent. The corresponding quantum operators are still anomalous (unless the Immirzi parameter takes the values $\gamma = \pm \sqrt{\zeta}$, where $\zeta$ denotes the internal signature, or $\gamma = \infty$). Therefore, in the new models (parts of) the simplicity constraints are implemented weakly to account for the anomaly. Also, the newly developed $\text{U}(N)$ tools [30–32] have recently been applied to solve the simplicity constraints [33–35].

In this paper, we are first going to take an unbiased look at them from the canonical perspective in the hope of finding new clues for how to implement the constraints correctly. Afterwards, we will compare our results to existing approaches from the spin foam literature and outline similarities and differences. We stress that we have not arrived at the conclusion that a certain kind of imposition will be the correct one, and thus, further research, centred around consistency considerations and the classical limit, has to be performed to find a satisfactory treatment for the simplicity constraints. Of course, in the end, by an experiment we have to decide which implementation, if any, will be the correct one. Since such experiments are missing up to now, the general guidelines are of course for the mathematical consistency of the approach, as well as for the comparison with the classical implementation of the simplicity constraints in $D = 3$, where the usual SU$(2)$ Ashtekar variables exist. If a satisfactory implementation in $D = 3$ can be constructed, the hope would then be that this procedure has a natural generalization to higher dimensions. Since parts of the very promising results developed from the spin foam literature are restricted to four dimensions, we will restrict ourselves to dimension-independent treatments in the main part of this paper.

The rest of this paper is divided into three parts. We begin with investigating the quadratic simplicity constraint operators which have been shown to be anomalous in [3]. It is illustrated that choosing a recoupling scheme for the intertwiner naturally leads to a maximal closing

\(^4\) See [7–9] for original literature works on gauge unfixing.
\(^5\) Here, ‘strongly’ means that the constraint operator annihilates physical states, $\hat{C} |\psi\rangle = 0 \forall |\psi\rangle \in \mathcal{H}_{\text{phys}}$. 

subset of simplicity constraint operators. Next, the solution to this subset is shown to allow for a natural unitary map to the SU(2)-based Ashtekar–Lewandowski Hilbert space in $D = 3$ and we finish the first part with several remarks on this quantization procedure. In section 3, we analyse the strong implementation of the linear simplicity constraint operators since they are non-anomalous from start. The resulting intertwiner space is shown to be one dimensional, which is problematic because this forbids the construction of a natural map to the SU(2)-based Ashtekar–Lewandowski Hilbert space. In contrast to the quadratic case, the linear simplicity constraint operators will be shown to be problematic when acting on edges. We discuss several possibilities of how to resolve these problems and finally introduce a mixed quantization, in which the linear simplicity constraints are substituted by the quadratic constraints plus a constraint ensuring the equality of the normals $N^I$ and $n^I(\pi)$. In section 4, we compare our results to the existing approaches from the spin foam literature. Finally, we give a critical evaluation of our results and conclude in section 5.

2. The quadratic simplicity constraint operators

2.1. A maximal closing subset of vertex constraints

In our companion papers [1–6], a canonical connection formulation of $(D + 1)$-dimensional Lorentzian general relativity was developed, using an SO$(D + 1)$-connection $A_{aIJ}$ and its conjugate momentum $\pi^{aIJ}$ as the canonical variables. Here, $a, b, \ldots = 1, \ldots, D$ are the spatial tensorial indices and $I, J, \ldots = 0, \ldots, D$ are the Lie algebra indices in the fundamental representation. A key input of the construction is the (quadratic) simplicity constraints

\[ \pi^{aIJ} \pi^{bIJ} \approx 0, \]  

which enforce, up to a topological sector present in $D = 3$, that $\pi^{aIJ} \approx 2n^IE_{aI}$, where $E_{aI}$ is an SO$(D + 1)$-valued vector density, a so-called hybrid vielbein, and $n^I$ is the unique (up to sign) normal defined by $n_IE_{aI} = 0$. Fixing the time gauge $n^I = (1, 0, \ldots, 0)$, one arrives at the ADM (extended phase space) formulation of general relativity with SO$(D)$ gauge invariance; see [2] for details. The second-class constraints which normally arise as stability conditions on the simplicity constraints are absent in our connection formulation, since they can be explicitly removed by the process of gauge unfixing after performing the Dirac analysis; see [2]. Essentially, they are gauge fixing conditions for the gauge transformations generated by the simplicity constraint, which change a certain part of the torsion of $A_{aIJ}$. The square of this part of the torsion is included in a respective decomposition of the Palatini action and thus results in the second-class partner for the simplicity constraint [2].

A quantization of the simplicity constraint using loop quantum gravity methods results in a complicated operator, since $\pi^{aIJ}$ becomes a flux operator which acts as the sum of all right-invariant vector fields associated with the different edges at a vertex. In order to facilitate the treatment of this quantum constraint, it has been shown in [3] that the necessary and sufficient building blocks of the quadratic simplicity constraint operator acting on a vertex $v$ are given by

\[ R_{[IJ}^\gamma R_{K]}^{\gamma f_{v'}} = 0 \quad \forall e, e' \in \{ e^{v''} \in E(\gamma) ; v = b(e'') \}, \]  

where $R_{[IJ}^{\gamma f_{v'}}$ is the right-invariant vector field associated with the edge $e$, $f_{v'}$ is the cylindrical function defined on an adapted graph $\gamma$, e.g., a spin network, $v$ is a vertex of $\gamma$, $E(\gamma)$ is the set of edges of $\gamma$ and $b(e)$ denotes the beginning of the edge $e$. The orientations of all edges are chosen such that they are outgoing of $v$. We note that these are exactly the off-diagonal simplicity constraints familiar from spin foam models; see, e.g., [11, 24].
Since not all of these building blocks commute with each other, i.e. the ones sharing exactly one edge, we have to resort to a non-standard procedure in order to avoid an anomaly in the quantum theory. The strong imposition of the above constraints, leading to the Barrett–Crane intertwiner [14], was discussed in [11]. A master constraint formulation of the vertex simplicity constraint operator was proposed in [3]; however, apart from providing a precise definition of the problem, this approach has not lead to a concrete solution up to now.

In this paper, we are going to explore a different strategy for implementing the quadratic vertex simplicity constraint operators which is guided by the following two natural requirements.

1. The imposition of the constraints should be non-anomalous.
2. The imposition of the simplicity constraint operator in $D = 3$ should, at least on the kinematical level, lead to the same Hilbert space as the quantization of the classical theory without a simplicity constraint. More precisely, there should exist a natural unitary map from the solution space of the quadratic simplicity constraint operators $H_{\text{simple}}$ to the Ashtekar–Lewandowski Hilbert space $H_{AL}$ in $D = 3$.

The concept of gauge unfixing [7–9] which was successfully used in order to derive the classical connection formulation of general relativity [1, 2] used in this paper was originally developed in the context of anomalous gauge theory, where it was observed that first-class constraints can turn into second-class constraints after quantization [36–40]. This is however precisely what is happening in our case: the classically Abelian simplicity constraints become a set of non-commuting operators due to the regularization procedure used for the fluxes. Thus, a natural question arises: How does a set of maximally commuting vertex simplicity constraint operators look like?

**Theorem 1.** Given a $N$-valent vertex $v \in \gamma$, the set

$$\epsilon_{IJKL}^\epsilon_{\epsilon_1 \epsilon_2} R_{\epsilon_1}^{IJ} R_{\epsilon_2}^{KL} = \cdots = \epsilon_{IJKL}^\epsilon_{\epsilon_N} R_{\epsilon_N}^{IJ} R_{\epsilon_N}^{KL} = 0$$  \hspace{1cm} (2.3)

generates a closed algebra of vertex simplicity constraint operators. Under the assumption that no linear combinations with different multi-indices $M = M_1 M_2 \cdots M_{D-3}$ are allowed,

the set is maximal in the sense that adding new vertex constraint operators spoils closure.

**Proof.** Closure can be checked by explicit calculation. In order to understand why the calculation works, recall that right-invariant vector fields generate the Lie algebra $so(D + 1)$ as [3]

$$[R_{IJ}^\epsilon, R_{KL}^\epsilon] = \frac{1}{2} \delta_{\epsilon \epsilon'} (\eta_{JK} R_{IL}^\epsilon + \eta_{IL} R_{JK}^\epsilon - \eta_{IK} R_{JL}^\epsilon - \eta_{JL} R_{IK}^\epsilon)$$  \hspace{1cm} (2.5)

and thus infinitesimal rotations. The commutativity of (2.3) has been discussed in [3]. Furthermore, we see that every element of (2.4) operates on (2.3) as an infinitesimal rotation.

\[\text{A superposition of different multi-indices seems to be highly unnatural since an anomaly with the Gauß constraint has to be expected. We are however currently not aware of a proof which excludes this possibility from the viewpoint of a maximal closing set.}\]
The same is also true for the elements in (2.4): taking the ordering from above, every constraint operates as an infinitesimal rotation on all constraints prior in the list. Since the commutator is antisymmetric in the exchange of its arguments, the closure, i.e. commutativity up to constraints, of (2.4) follows.

To prove the maximality of the set we show that, having chosen a subset of simplicity constraints as given in (2.3) and (2.4), adding any other linear combination of the building blocks (2.2) spoils the closure of the algebra. To this end, we make the most general Ansatz

\[
\sum_{1 \leq i < j < N} \alpha_{ij} \epsilon_{IJKL} R^{ij}_{KL} \equiv 0
\]

for an \( N \)-valent vertex. Note that the diagonal terms \((i = j)\) are proportional to (2.3) and therefore do not have to be taken into account in the above sum and that \( R_N = \sum_{i=1}^{N-1} R_i \) can be dropped due to gauge invariance. Moreover, \( \alpha_{ij} \) can be chosen such that for fixed \( j \) not all \( \alpha_{ij} \) \((i < j)\) are equal. Otherwise, with \( \alpha_{ij} := \alpha_j \), we find the term \( \alpha_j \epsilon_{IJKL} R^{ij}_{KL} \) in the sum, which can be expressed as a linear combination of (2.3) and (2.4) and therefore can be dropped. Consider

\[
\left[ \epsilon_{IJKL} R^{ij}_{KL}, \epsilon_{ABCDEF} (\alpha_{13} R^{12}_{CD} + \alpha_{23} R^{AB}_{CD} + \cdots) \right]
\]

\[
\approx \sum_{j=3}^{N-1} 2 \alpha_{1j} \epsilon_{IJKL} R^{ij}_{KL} f^{KL AB}_{MN} R^{MN}_{1} \epsilon_{ABCDEF} R^{CD}_{j}
\]

\[
+ \sum_{j=3}^{N-1} 2 \alpha_{2j} \epsilon_{IJKL} R^{ij}_{KL} f^{KL AB}_{MN} R^{MN}_{2} \epsilon_{ABCDEF} R^{CD}_{j}
\]

\[
\approx \sum_{j=3}^{N-1} 2 (\alpha_{1j} - \alpha_{2j}) \epsilon_{IJKL} R^{ij}_{KL} f^{KL AB}_{MN} R^{MN}_{3} \epsilon_{ABCDEF} R^{CD}_{j},
\]

where we dropped terms proportional to (2.3) in the first step and in the second step. For a closing algebra, the right-hand side of (2.7) necessarily has to be proportional to (a linear combination of) simplicity building blocks (2.2). Terms containing \( R_j \) \((j \geq 3)\) have to vanish separately. (In general, one could make use of gauge invariance to ‘mix’ the contributions of different \( R_j \). However, in the case at hand this will produce terms containing \( R_N \), which do not vanish if the contributions of different \( R_j \)’s did not already vanish separately.)

We start with the case \( D = 3 \). The summands on the right-hand sides of (2.7) are proportional to

\[
\delta_{ABC} (R_j)^{AB} (R_2)^{ij} (R_1)^K, 
\]

where we used the notation \( \delta^{lij}_{kl} := n! \delta_{lij} \delta_{jkl} \). To show that this expression cannot be rewritten as a linear combination of the building blocks (2.2), we antisymmetrize the indices \([ABI], [ABK] \) and \([IJK] \) and find in each case that the result is zero.

For \( D > 3 \), the summands are proportional to

\[
\delta_{ABC} (R_j)^{ABC} (R_2)^{ijkl} (R_1)^K, 
\]

Whatever multi-index \( \overline{E} \) we might have chosen in the Ansatz (2.6), we can always restrict attention to those simplicity constraints in the maximal set which have the same multi-index \( \overline{M} = \overline{E} \). Then, the same calculation as in the case of \( D = 3 \) shows that the antisymmetrizations of the indices \([ABI], [ABK] \) and \([IJK] \) vanish.

Therefore, the only possibilities are (a) the trivial solution \( \alpha_{1j} = \alpha_{2j} = 0 \) and (b) \( \alpha_{1j} = \alpha_{2j}(\neq 0) \), which implies that the terms on the right-hand side of (2.7) are a rotated version of \( \epsilon_{IJKL} R^{ij}_{KL} \). The second option (b) is, for \( j = 3 \), excluded by our choice of \( \alpha_{ij} \) and
Figure 1. Recoupling scheme corresponding to the subset of quadratic vertex simplicity constraint operators (2.4).

we must have $\alpha_{13} = \alpha_{23} = 0$. Next, consider $j = 4$ and suppose we have $\alpha_{14} = \alpha_{24} = \alpha' \neq 0$. Then, we can define $\alpha'_{34} := \alpha_{34} - \alpha'$ and find the terms $\alpha'_{\epsilon IJKL} R_{e_1}^{KL} R_{e_2}^{KL} + \alpha'_{\epsilon IJKL} R_{e_1}^{KL} R_{e_2}^{KL}$ in (2.6). The first term again is already in the chosen set, which implies that we can set $\alpha_{14} = \alpha_{24} = 0$ w.l.o.g. by changing $\alpha_{34} \rightarrow \alpha'_{34}$. (We will drop the prime in the following.)

This immediately generalizes to $j > 4$, and we have w.l.o.g. $\alpha_{ij} = 0 (3 \leq j < N)$.

Suppose we have calculated the commutators of $\epsilon_{\epsilon IJKL} R_{e_1}^{KL} R_{e_2}^{KL} (i = 2, \ldots, n)$ with (2.6) and found that for closure, we need $\alpha_{ij} = 0$ for $1 \leq i \leq n$ and $i < j < N$. Then,

$$\left[ \epsilon_{\epsilon IJKL} R_{e_1}^{KL} R_{e_2}^{KL}, \epsilon_{\epsilon ABCD} R_{e_1}^{CD} \right] \approx \sum_{j=(n+2)}^{N-1} 2\alpha_{(n+1)j} \epsilon_{\epsilon IJKL} R_{e_1}^{KL} R_{e_2}^{KL} R_{e_1}^{KL} R_{e_2}^{KL} = 0$$

which, by the reasoning above, again is not a linear combination of any simplicity building blocks for any choice of $\alpha_{(n+1)j}$, and therefore, only the trivial solution $\alpha_{(n+1)j} = 0 (n + 1 < j < N)$ leads to closure of the algebra. □

2.2. The solution space of the maximal closing subset

In order to interpret this set of constraints, recall from [3] that the constraints in (2.3) are the same as the diagonal simplicity constraints acting on edges of $\gamma$ and can be solved by demanding the edge representations to be simple. The remaining constraints (2.4) can be interpreted as specifying a recoupling scheme for the intertwiner $\iota$ at $v$: couple the representations on $e_1$ and $e_2$, then couple this representation to $e_3$, and so forth, see figure 1. We call the intermediate virtual edges $e_{12}$, $e_{123}$, ... and denote the highest weights of the representations thereon by $\lambda_{12}, \lambda_{123}, \ldots$. Since we can use gauge invariance at all the intermediate intertwiners in the recoupling scheme, e.g., $R_{e_1} + R_{e_2} = R_{e_{12}}$, we have

$$\epsilon_{\epsilon IJKL} R_{e_1}^{KL} R_{e_2}^{KL} = 0$$

and thus that the representation on $e_{12}$ has to be simple, i.e.

$$\lambda_{12} = (\lambda_{12}, 0, \ldots, 0) \quad \lambda_{12} = 0, 1, 2, \ldots$$
Using the same procedure, all intermediate representations are required to be simple and the intertwiner is labelled by \( N - 3 \) ‘spins’ \( \lambda_i \in \mathbb{N}_0 \). We call an intertwiner where all internal lines are labelled with simple representations simple.

Denote by \( \mathcal{I}_{\text{SU}(2)}^N \) the set of \( N \)-valent SU(2) intertwiners and by \( \mathcal{I}_{\text{Spin}(D+1)}^N \) the set of simple \( N \)-valent spin \( (D + 1) \) intertwiners. Recalling that an \( N \)-valent SU(2) intertwiner can be expressed in the same recoupling basis and calling the intermediate spins \( j_i \), we see that the map

\[
F : \mathcal{I}_{\text{Spin}(D+1)}^N \rightarrow \mathcal{I}_{\text{SU}(2)}^N
\]

\( \frac{1}{2} \lambda_i \mapsto j_i \)

is unitary (with respect the scalar products induced by the respective Ashtekar–Lewandowski measures, see [3]). The motivation for the factor \( 1/2 \) comes from the fact that \( \vec{\lambda} = (1, 0) \) in \( D = 3 \) corresponds to the familiar \( j_+ = j_- = 1/2 \) and the area spacings of the SO(4)-based and the SU(2)-based theories agree using this identification, cf [3].

2.3. Remarks

(1) Since the choice of the maximal closing subset of the simplicity constraint operators is arbitrary, no recoupling basis is preferred \textit{a priori}. On the SU(2) level, a change in the recoupling scheme amounts to a change of basis in the intertwiner space and therefore poses no problems. On the level of simple spin \( (D + 1) \) representations, however, a choice in the recoupling scheme affects the property ‘simple’, since the non-commutativity of constraint operators belonging to different recoupling schemes means that kinematical states cannot have the property simple in both schemes.

(2) There exist recoupling schemes which are not included in the above procedure, e.g., take \( N = 6 \) and the constraints \( \epsilon R_{12} R_{12} = \epsilon R_{34} R_{34} = \epsilon R_{56} R_{56} = 0 \) and couple the three resulting simple representations. The theorem should however generalize to those additional recoupling schemes.

(3) It is doubtful if the action of the Hamiltonian constraint leaves the space of simple intertwiners in a certain recoupling scheme invariant. To avoid this problem, one could use a projector on the space of simple intertwiners in a certain recoupling scheme to restrict the Hamiltonian constraint on this subspace and average later on over the different recoupling schemes if they turn out to yield different results. The possible drawbacks of such a procedure are however currently unclear to the authors and we refer to further research. The construction of such a projector can be seen as a quantum analogue of the gauge unfixing process familiar from our companion paper [2]. A possible strategy to find a Hamiltonian constraint operator which leaves the solution space of a first-class subset invariant is to construct a gauge unfixing projector which adds vertex simplicity constraints which are not in the first-class subset to the Hamiltonian constraint such that it commutes with the first-class subset.

(4) It would be interesting to check whether the dropped constraints are automatically solved in the weak operator topology (matrix elements with respect to solutions to the maximal subset).

(5) The imposition of the constraints can be stated as the search for the joint kernel of a maximal set of commuting generalized area operators

\[
\text{Ar}_{\Sigma}[\sigma] := \sum_{U \in \mathcal{U}} \sqrt{\frac{1}{2}} \epsilon_{ijkl} \pi^{ij}(S_U) \pi^{kl}(S_U).
\]

Notice, however, that for \( D > 3 \) these generalized area operators, just as the simplicity constraints, are not gauge invariant, while in \( D = 3 \) they are gauge invariant.
(6) In $D = 3$, we have the following special situation.

We have two classically equivalent extensions of the ADM phase at our disposal whose respective symplectic reduction reproduces the ADM phase space. One of them is the Ashtekar–Barbero–Immirzi connection formulation in terms of the gauge group $SU(2)$ with additional $SU(2)$ Gauß constraint next to spatial diffeomorphism and Hamiltonian constraint, and the other is our connection formulation in terms of $SO(4)$ with additional $SO(4)$ Gauß constraint and simplicity constraint. Both formulations are completely equivalent, and thus, one should expect that also the quantum theories are equivalent in the sense that they have the same semiclassical limit. Let us ask a stronger condition, namely that the joint kernel of $SO(4)$ Gauß and simplicity constraint of the $SO(4)$ theory is unitarily equivalent to the kernel of the $SU(2)$ Gauß constraint of the $SU(2)$ theory. To investigate this first from the classical perspective, we split the $SO(4)$ connection and its conjugate momentum $(A^{ij}, \pi_{ij})$ into self-dual and anti-self-dual parts $A^\pm(A, \pi^\pm)$ which then turn out to be conjugate pairs again. It is easy to see that the $SO(4)$ Gauß constraint $G_{ij}$ splits into two $SU(2)$ Gauß constraints $G^\pm_j$, one involving only self-dual variables and the other only anti-self-dual ones which therefore mutually commute as one would expect. The $SO(4)$ Gauß constraint now asks for separate $SU(2)$ gauge invariance for these two sectors. Thus, a quantization in the Ashtekar–Isham–Lewandowski representation would yield a kinematical Hilbert space with an orthonormal basis $T^+_k \otimes T^-_k$, where $S_k$ are usual $SU(2)$-invariant spin networks. The simplicity constraint, which in $D = 3$ is Gauß invariant and can be imposed after solving the Gauß constraint, from a classical perspective suggests that the double-density inverse metrics $q^{ab}_\pm = \pi^{\alpha\beta}_\pm \delta^{ab}_\pm$ are identical. This is classically equivalent to the statement that the corresponding area functions $A_{+}(S)$ are identical for every $S$. The corresponding statement in the quantum theory is, however, again anomalous because it is well known that area operators do not commute with each other. On the other hand, neglecting this complication for a moment, it is clear that the quantum constraint can only be satisfied on vectors of the form $T^+_k \otimes T^-_k$ for all $S$ if $s_+, s_- \in \{1, -1\}$ share the same graph and $SU(2)$ representations on the edges because if $S$ cuts a single edge transversally, then the area operator is diagonal with an eigenvalue $\sqrt{f(j \pm T)}$ and we can always arrange such an intersection situation by choosing suitable $S$. By a similar argument, one can show that the intertwiners at the edges have to be the same. But this is only a sufficient condition because in a sense there are too many quantum simplicity constraints due to the anomaly. However, the discussion suggests that the joint kernel of both $SO(4)$ and simplicity constraint is the closed linear span of vectors of the form $T^+_k \otimes T^-_k$ for the same spin network $s = s_+ = s_-$. The desired unitary map between the Hilbert spaces would therefore simply be $T_k \mapsto T^+_k \otimes T^-_k$.

This can be justified abstractly as follows. From all possible area operators, pick a maximal commuting subset $A^\pm_k$ using the axiom of choice (i.e. pick a corresponding maximal set of surfaces $S_k$). We may construct an adapted orthonormal basis $T^\pm_k$ diagonalizing all of them, such that $A^\pm_k T^\pm_k = \lambda_k T^\pm_k$. Now the constraint

$$A^+_k \otimes 1 = 1 \otimes A^-_k$$

can be solved on vectors $T^+_k \otimes T^-_k$ by demanding $\lambda_+ = \lambda_-$. The desired unitary map would then be $T_k \mapsto T^+_k \otimes T^-_k$. Thus, the question boils down to asking whether a maximal closing subset can be chosen such that the eigenvalues $\lambda$ are just the spin networks $s$. We leave this to future research.

If the maximal set still separates the points of the classical configurations space, this should leave no room for degeneracies, i.e. $\lambda_k$ completely specify the eigenvector. We will assume this to be the case for the following argument.
In $D \neq 3$, the aforementioned split into self-dual and anti-self-dual sector is meaningless and we must stick with the dimension-independent scheme outlined above. An astonishing feature of this scheme is that after the proposed implementation of the simplicity constraints, the size of the kinematical Hilbert space is the same for all dimensions $D \geqslant 3$. By ‘size’, we mean that the spin networks are labelled by the same sets of quantum numbers on the graphs. Of course, before imposing the spatial diffeomorphism constraint, these graphs are embedded into spatial slices of different dimension and thus provide different amounts of degrees of freedom. However, after the implementation of the diffeomorphism constraint, most of the embedding information will be lost and the graphs can be treated almost as abstract combinatorial objects. Let us neglect here, for the sake of the argument, the possibility of certain remaining moduli, depending on the amount of diffeomorphism invariance that one imposes, which could a priori be different in different dimensions. In the case that the proposed quantization would turn out to be correct, i.e. allow for the correct semiclassical limit, this would mean that the dimensionality of space would be an emergent concept dictated by the choice of semiclassical states which provide the necessary embedding information. A possible caveat to this argument is the remaining Hamiltonian constraint and the algebra of Dirac observables which critically depend on the dimension (for instance through the volume operator or dimension-dependent coefficients, see [1, 2]) and which could require to delete different amounts of degrees of freedom depending on the dimension.

This idea of dimension emergence is not new in the field of quantum gravity; however, it is interesting to possibly see here a concrete technical realization which appears to be forced on us by demanding anomaly freedom of the simplicity constraint operators. Of course, these speculations should be taken with great care: the number of degrees of freedom of the classical theory certainly does strongly depend on the dimension, and therefore, the speculation of dimension emergence could fail exactly when we try to construct the semiclassical sector with the solutions to the simplicity constraints advertised above. This would mean that our scheme is wrong. On the other hand, there are indications [41] that the semiclassical sector of the LQG Hilbert space already in $D = 3$ is entirely described in terms of six-valent vertices. Therefore, the higher valent graphs, which in $D = 3$ could correspond to pure quantum degrees of freedom, could account for the semiclassical degrees of freedom of higher dimensional general relativity. Since there is no upper limit to the valence of a graph, this would mean that already the $D = 3$ theory contains all higher dimensional theories!

Obviously, this puzzle asks for thorough investigation in future research.

The discussion reveals that we should compare the amount of degrees of freedom that the classical and the quantum simplicity constraint removes. This is a difficult subject, because there is no well-defined scheme that attributes quantum to classical degrees of freedom unless the Hilbert space takes the form of a tensor product, where each factor corresponds to precisely one of the classical configuration degrees of freedom. The following ‘counting’ therefore is highly heuristic and speculative.

In the case $D = 3$, the classical simplicity constraints remove six degrees of freedom from the constraint surface per point on the spatial slice. In order to count the quantum degrees of freedom that are removed by the quantum simplicity constraint when acting on a spin network function, we make the following admittedly naïve analogy.

We attribute to a point on the spatial slice an $N$-valent vertex $v$ of the underlying graph $\gamma$ which is attributed to the spatial slice. This point is equipped with degrees of freedom labelled by edge representations and the intertwiner. Every edge incident at $v$ is shared
by exactly one other vertex (or returns to \( v \) which however does not change the result).

Therefore, only half of the degrees of freedom of an edge can be attributed to one vertex. We take as edge degrees of freedom the \( \left\lceil \frac{D+1}{2} \right\rceil \) Casimir eigenvalues of \( SO(D+1) \) labelling the irreducible representation. The edge simplicity constraint removes all but one of these Casimir eigenvalues; thus, per edge \( \left\lceil \frac{D+1}{2} \right\rceil \) edge degrees of freedom are removed. Furthermore, a gauge-invariant intertwiner is labelled by a recoupling scheme involving \( N-3 \) irreducible representations not fixed by the irreducible representations carried by the edges adjacent to the vertex in question, which are fully attributed to the vertex (there are \( N-2 \) virtual edges coming from coupling 1, 2, 3, etc, until \( N \) but the last one is fixed due to gauge invariance). We take as vertex degrees of freedom these \( N-3 \) irreducible representations each of which is labelled again by \( \left\lceil \frac{D+1}{2} \right\rceil \) Casimir eigenvalues.

The vertex simplicity constraint again deletes all but one of these eigenvalues; thus, it removes \( (N-3)\left\lceil \frac{D+1}{2} \right\rceil \) quantum degrees of freedom. We conclude that the quantum simplicity constraint removes

\[
\left( N - 3 + \frac{N}{2} \right) \left\lceil \frac{D-1}{2} \right\rceil \]

quantum degrees of freedom per point (\( N \)-valent vertex), where \( N-3 \) accounts for the vertex and \( N/2 \) for the \( N \) edges counted with half weight as argued above. This is to be compared with the classical simplicity constraint which removes \( D^2(D-1)/2 - D \) degrees of freedom per point. Requiring equality we see that vertices of a definitive valence \( N_D \) are preferred in \( D \) spatial dimensions which for large \( D \) grows quadratically with \( D \). Specifically, for \( D = 3 \), we find \( N_3 = 6 \).

Thus, our na"ive counting astonishingly yields the same preference for six-valent graphs in \( D = 3 \) as has been obtained in [41] by completely different methods. From the analysis of [41], it transpires that \( N_3 = 6 \) has an entirely geometric origin and one thus would rather expect \( N_D = 2D \) (hypercubulations) and this may indicate that our counting is incorrect.

### 3. The linear simplicity constraint operators

#### 3.1. Regularization and anomaly freedom

In [5], the connection formulation sketched at the beginning of the previous section was altered in that it contains linear simplicity constraints \( \epsilon_{IJKL}N^d \pi^{dIJK} \approx 0 \) and an independent normal \( N^d \) as phase space variables. The normal Poisson commutes with both the connection \( A_{dI} \) and its momentum \( \pi^{dI} \) and has its own canonical momentum \( P_I \). The necessity for this independent normal did not stem from the anomaly encountered when looking at the quadratic quantum simplicity constraints, but from the observation that it was needed to extend the connection formulation to higher dimensional supergravities.

Since the linear simplicity constraint is a vector of density weight one, it is most naturally smeared over \( (D-1) \)-dimensional surfaces. The regularization of the objects

\[
S^b(S) := \int_S h^{IJKL}(x)\epsilon_{IJKL}N^d(x)\pi^{dIJK}(x)\epsilon_{ab...h}\;dx^b \wedge \cdots \wedge dx^{b-1},
\]

where \( S^b \) denotes the linear simplicity constraint, \( S \) denotes a \( D-1 \)-surface and \( h^{IJKL} \) denotes an arbitrary semianalytic smearing function of compact support and, therefore, is completely analogous to the case of flux vector fields. The corresponding quantum operator

\[
\hat{S}^b(S)f = \hat{\chi}^{b}\hat{\epsilon}^b(S)f = p_{\gamma} p_{\gamma} \sum_{e \in \gamma} \epsilon(e, S)\epsilon_{IJKL}b^{IJKL}(b(e))\hat{N}^d(b(e))R^k x^k f_\gamma
\]

(3.2)
has to annihilate physical states for all surfaces \( S \subset \sigma \) and all semianalytic functions \( h/\mathcal{M} \) of compact support, where \( p_\gamma \) denotes the cylindrical projection and \( \gamma_S \) is a graph adapted to the surface \( S \). Since we can always choose surfaces which intersect a given graph only at one point, this implies that the constraint has to vanish when acting on single points of a given graph. In [3], it has been shown that the right-invariant vector fields actually are in the linear span of the flux vector fields. Therefore, it is necessary and sufficient to demand that

\[
\epsilon_{ijklM}h^{ijM}(b(e))\tilde{N}^l(b(e))R^k_{ce} \cdot f_p = 0
\]

for all points of \( \gamma \) (which can be seen as the beginning point of edges by suitably subdividing and inverting edges). Since \( \tilde{N}^l \) acts by multiplication and commutes with the right-invariant vector fields, see [5] for details, the condition is equivalent to

\[
\tilde{R}_e^c \cdot f_p = 0,
\]

i.e. the generators of rotations stabilizing \( N^l \) have to annihilate physical states. Before imposing these conditions on the quantum states, we have to consider the possibility of an anomaly. Classically and before using the singular smearing of holonomies and fluxes, both the linear and the quadratic simplicity constraints are Poisson self-commuting. The quadratic constraint is known to be anomalous both in the spin foam [24] as well as in the canonical picture [25, 3] and thus should not be imposed strongly. Also, the linear simplicity constraint is anomalous when using a non-zero Immirzi parameter. (At least if \( \gamma \neq 1 \) in the Euclidean theory. But \( \gamma = 1 \) is ill-defined for \( \text{SO}(4) \), see, e.g., [42].) Surprisingly, in the case at hand and without an Immirzi parameter in four dimensions, we do not find an anomaly. However, that is just because the generators of rotations stabilizing \( N^l \) form a closed subalgebra! Direct calculation yields, choosing (without loss of generality) \( \gamma_{SS} \) to be a graph adapted to both surfaces \( S, S' \),

\[
\left[ \tilde{S}^l_{\gamma r} (S), \tilde{S}^l_{\gamma r} (S') \right]_{f_{\gamma r}} = \left[ \sum_{e \in \gamma_r} \tilde{R}^l_{e} \right]_{f_{\gamma r}} = \sum_{e \in \gamma_r} \tilde{R}^l_{e},
\]

where the operator in the last line is in the linear span of the vector fields \( \tilde{S}^l (S) \). The classical constraint algebra is not reproduced exactly (the commutator does not vanish identically), but the algebra of quantum simplicity constraints closes; they are of the first class. Therefore, strong imposition of the quantum constraints does make mathematical sense.

Note that up to now, we did not solve the Gauß constraint. The quantum constraint algebra of the simplicity and the Gauß constraint can easily be calculated and reproduces the classical result

\[
\left[ \tilde{S}^l (S), \tilde{G}^{AB}[\Lambda_{AB}] \right]_{f_{\gamma r}} = \sum_{e \in \gamma_r} \epsilon(e',S)\epsilon_{ijklM}b_{ijklM}(v)\tilde{N}^l(v)R^k_{ce},
\]

\[
\Lambda_{AB}(v) \left( \sum_{e \in \gamma_r} R^c_{ce} + R^c_{ce} \right)_{f_{\gamma r}}
\]

\footnote{Use the decomposition of \( X_M \) into its rotational \( \tilde{X}_{M} := \tilde{\eta}_{M}X_{M} \) and ‘boost’ parts \( \tilde{X}_{N} := -\xi N^JX_{M} \) with respect to \( N^J \) in (3.3), where \( \tilde{\eta}_{M} = \eta_{M} - \xi N^Jx_{M} \) and \( \xi = 1 \) for \( \text{SO}(D + 1) \) and \( \xi = -1 \) for \( \text{SO}(1, D) \) as internal gauge groups. It follows that \( \tilde{\eta}_{M}N_J = 0 \).}
where we used $R^A_B := \frac{1}{2} \left( N^A \frac{\partial}{\partial N^B} - N^B \frac{\partial}{\partial N^A} \right)$. It follows that the simplicity constraint operator does not preserve the Gauß-invariant subspace (in other words, as in the classical theory, the Gauß constraint does not generate an ideal in the constraint algebra). This implies that the joint kernel of both Gauß and simplicity constraint must be a proper subspace of the Gauß-invariant subspace. It is therefore most convenient to look for the joint kernel on in the kinematical (non-Gauß-invariant) Hilbert space.

### 3.2. Solution on the vertices

Consider a slight modification of the usual gauge-variant spin network functions, where the intertwiners $i_v = i_v(N)$ are the square integrable functions of $N^I$. Let $v$ be a vertex of $\gamma$ and $e_1, \ldots, e_n$ be the edges of $\gamma$ incident at $v$, where all orientations are chosen such that the edges are all outgoing at $v$. Then, we can write the modified spin network functions

$$T_{\gamma,\vec{\gamma},i}(A, N) := (i_v(N))_{\vec{K}_v} \prod_{j=1}^n (\pi_{i_v}(h_v(A)))_{\vec{K}_j} \hat{M}_v, \vec{K}_v \cdot \vec{K}_i,$$

where $\hat{M}_v$ contracts the indices corresponding to the endpoints of the edges $e_i$ and represents the rest of the graph $\gamma$. These states span the combined Hilbert space for the normal field and the connection $H_T = H_{\text{grav}} \otimes H_N$ (cf [5]) and they will prove convenient for solving the simplicity constraints. Choose the surface $S'$ such that it intersects a given graph $\gamma'$ only in the vertex $v' \in \gamma'$. The action of $\hat{S}'(S')$ on the vertex $v'$ of a spin network $T_{\gamma',\vec{\gamma}',i}(A, N)$ implies with (3.4) that

$$\hat{S}'(S')_{v'} T_{\gamma',\vec{\gamma}',i}(A, N) = 0 \iff \text{tr} \left( (i_v(N) \tau_{\gamma'}^I) \otimes_{i=1}^n \pi_{i_v}(h_v(A)) \cdot M_v \right) = 0 \text{ for all } v',$$

where $\tau_{\gamma'}^I$ denote the generators of $SO(D + 1)$ in the representation $\pi_{i_v}$ of the edge $v$ and the bar again denotes the restriction to rotational components (w.r.t. $N^I$). The above equation implies that the intertwiner $i_v$, seen as a vector transforming in the representation $\pi_{i_v}$ dual to $\pi_{e_i}$ of the edge $e$, has to be invariant under the $SO(D)_N$ subgroup which stabilizes $N^I$. By definition [43], the only representations of $SO(D + 1)$ which have in their space nonzero vectors which are invariant under a $SO(D)$ subgroup are of the representations of class one (cf also the appendix), and they exactly coincide with the simple representations used in spin foams [11]. It is easy to see that the dual representations (in the sense of group theory) of simple representations are simple representations. Therefore, all edges must be labelled by simple representations of $SO(D + 1)$. Moreover, $SO(D)$ is a massive subgroup of $SO(D + 1)$ [43], so that the (unit) invariant vector $\xi_{i_v}(N)$ in the representation $\pi_{i_v}$ is unique, which implies that the allowed intertwiners $i_v(N)$ are given by the tensor product of the invariant vectors of all $n$ edges and potentially an additional square integrable function $F_i(N)$, $i_v(N) = \xi_{i_v}(N) \otimes \cdots \otimes \xi_{i_v}(N) \otimes F_i(N)$. Going over to normalized gauge-invariant spin network...
functions implies that $F_c(N) = 1$, and the resulting intertwiner space solving the simplicity and Gauß constraint becomes one dimensional, spanned by $I_c(N) := \xi_{i_1}(N) \otimes \cdots \otimes \xi_{i_n}(N)$. We will call these intertwiners and vertices coloured by them linear simple. For an instructive example of the linear-simple intertwiners, consider the defining representation (which is simple since the highest weight vector is $\Lambda = (1, 0, \ldots, 0)$, cf the appendix). The unit vector invariant under rotations (w.r.t. $N^\alpha$) is given by $N^\alpha$ and for edges in the defining representation incoming at $v$ we simply contract $h^\alpha_I N_{ij}$. If the constraint is acting on an interior point of an analytic edge, this point can be considered as a trivial two-valent vertex and the above result applies. Since this has to be true for all surfaces, a spin network function solving the constraint would need to have linear-simple intertwiners at every point of its graph $\gamma$, i.e. at infinitely many points, which is in conflict with the definition of cylindrical functions (cf [44]). In the next section, we comment on a possibility of how to implement this idea.

### 3.3. Edge constraints

As noted above, the imposition of the linear simplicity constraint operators acting on edges is problematic, because it does not, as one might have expected, single out simple representations, but demand that at every point, where it acts, there should be a linear-simple intertwiner. The problem with this type of solution is that all intertwiners, even trivial intertwiners at all interior points of edges, have to be linear simple, which is however in conflict with the definition of a cylindrical function; in other words, there would be no holonomies left in a spin network because every point would be an $N$-dependent vertex.

It could be possible to resolve this issue using a rigging map construction [45–47] of the type

$$
\eta(T_{p, p_{\gamma}, \gamma}) = \lim_{\Psi_p \ni \Psi_p \to \infty} C(p_{\gamma}, T_{p, T_{p_{\gamma}}, T_{p_{\gamma}, T_{p_{\gamma}}}}) \langle T_{p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}}}} \rangle_{\text{kin}},
$$

(3.9)

where $\Psi_{\gamma}$ is the set of finite point sets $p$ of a graph $\gamma$, $p = \{x_i\}_{i=1}^N | x_i \in \gamma \forall i, N < \infty \}$. $\Psi_\gamma$ is partially ordered by inclusion, $q \supset p$ if $p$ is a subset of $q$, so that the limit is meant in the sense of net convergence with respect to $\Psi_{\gamma}$. By the prescription $T_{p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}}}}$, we mean the projection of $T_{p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}}}}$ onto linear-simple intertwiners at every point in $p$ and $C(p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}}})$ is a numerical factor. Assuming this to work, consider any surface $S$ intersecting $\gamma$. We (heuristically) find

$$
\eta(T_{p_{\gamma}}|\hat{S}^\beta(S)T_{p_{\gamma}}) = \lim_{\Psi_p \ni \Psi_p \to \infty} C(p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}}}})|\hat{S}^\beta(S)T_{p_{\gamma}}\rangle_{\text{kin}} = \lim_{\Psi_p \ni \Psi_p \to \infty} C(p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}}})|[\hat{S}^\beta(S)]^\dagger T_{p_{\gamma}}\rangle_{\text{kin}} = \lim_{\Psi_p \ni \Psi_p \to \infty} C(p_{\gamma}, T_{p_{\gamma}, T_{p_{\gamma}}})\langle \hat{S}^\beta(S)T_{p_{\gamma}}| T_{p_{\gamma}}\rangle_{\text{kin}} = 0,
$$

(3.10)

since the intersection points of $S$ with $\gamma$ will eventually be in $p_{\gamma}$ and $\hat{S}^\beta(S)$ is self-adjoint.

We were however not able to find such a rigging map with satisfactory properties. It is especially difficult to handle observables with respect to the linear simplicity constraint and to implement the requirement that the rigging map has to commute with observables. It therefore seems plausible to look for non-standard quantization schemes for the linear simplicity constraint operators, at least when acting on edges. Comparison with the quadratic simplicity constraint suggests that also the linear constraint should enforce simple representations on the edges; see the following remarks and section 3.5 for ideas on how to reach this goal.
3.4. Remarks

The intertwiner space at each vertex is one dimensional, and thus, the strong solution of the unaltered linear simplicity constraint operator contrasts the quantization of the classically imposed simplicity constraint at first sight. A few remarks are appropriate.

(1) One could argue that the intertwiner space at a vertex $v$ is infinite dimensional by taking into account holonomies along edges $e'$ originating at $v$ and ending in a one-valent vertex $v'$. Since $e'$ and $v'$ are assigned in a unique fashion to $v$ if the valence of $v$ is at least 2, we can consider the set $\{v, e', v'\}$ as a new ‘non-local’ intertwiner. Since we can label $e'$ with an arbitrary simple representation, we obtain an infinite set of intertwiners which are orthogonal in the above scalar product. This interpretation however does not mimic the classical imposition of the simplicity constraints or the above imposition of the quadratic simplicity constraint operators.

(2) The main difference between the formulations of the theory with quadratic and linear simplicity constraints, respectively, is the appearance of the additional normal field sector in the linear case. Thus, one could expect that one would recover the quadratic simplicity constraint formulation by ad hoc averaging the solutions of the linear constraint over the normal field dependence with the probability measure $\nu_N$ defined in [5]. Indeed, if one does so, then one recovers the solutions to the quadratic simplicity constraints in terms of the Barrett–Crane intertwiners in $D = 3$ and higher dimensional analogues thereof as has been shown long ago by Freidel et al [11]. A similar observation has been made in [48]. Such an average also deletes the solutions with ‘open ends’ of the previous remark by an appeal to Schur’s lemma. Since after such an average the $N$ dependence of all solutions disappears, we can drop the $\mu_N$ integral in the kinematical inner product since $\mu_N$ is a probability measure. The resulting effective physical scalar product would then be the Ashtekar–Lewandowski scalar product of the theory between the solutions to the quadratic simplicity constraints. Such an averaging would also help with the solution of the edge constraints, since a two-valent linear-simple intertwiner is averaged as

$$\int_{SO} d\nu(N) \tilde{\xi}_i^\alpha(N) \tilde{\xi}_j^\beta(N) = \frac{1}{d\pi} \delta^{\alpha\beta},$$

(3.11)

thus yielding a projector on simple representations.

(3) It can be easily checked that the volume operator as defined in [3], and therefore, also more general operators, like the Hamiltonian constraint, do not leave the solution space to the linear (vertex) simplicity constraints invariant. A possible cure would be to introduce a projector $\mathcal{P}_S$ on the solution space and redefine the volume operator as $\hat{V} : = \mathcal{P}_S \hat{V} \mathcal{P}_S$. Such procedures are however questionable on the general ground that anomalies can always be removed by projectors.

(4) If one accepts the usage of the projector $\mathcal{P}_S$, calculations involving the volume operator simplify tremendously since the intertwiner space is one dimensional. We will give a few examples which can be calculated by hand in a few lines, restricting ourselves to the defining representation of $SO(D + 1)$, where the $SO(D)_N$-invariant unit vector is given by $N^I$.

Having direct access to $N^I$, one can base the quantization of the volume operator on the classical expression

$$\det q = \left| \frac{1}{D!} \epsilon_{I_1 \cdots I_D} N^I (\pi^{a_1 K_1} N_{K_1}) \cdots (\pi^{a_D K D} N_{K D}) \epsilon_{a_1 \cdots a_D} \right|^\frac{1}{2}.$$  

(3.12)
In the case $D + 1$ is uneven, this choice is much easier than the expression quantized in [3]. In the case $D + 1$ is even, the above choice is of the same complexity as the one in [3], but leads to a formula applicable in any dimension and therefore, for us, is favoured. Proceeding as in [3], we obtain for the volume operator

\[
\hat{\mathcal{V}}(R) = \int_R d^D p \delta(p)\gamma, \quad \hat{\mathcal{V}}(p) = \left(\frac{\hbar}{2}\right)^{\frac{D}{2}} \sum_{v \in V(\gamma)} \delta^D(p, v)\hat{V}_{v, \gamma},
\]

\[
\hat{V}_{v, \gamma} = \left| \frac{\hbar}{D!} \sum_{e_1, \ldots, e_D} s(e_1, \ldots, e_D)q_{e_1, \ldots, e_D} \hat{\mathcal{P}}_{\gamma} \right|^{\frac{1}{2}},
\]

\[
\hat{q}_{e_1, \ldots, e_D} = \epsilon_{IJ, \ldots, D} \hat{N}^{\delta(1)}(\hat{N}_{e_1}^{\delta(1)}) \cdots (\hat{N}_{e_D}^{\delta(1)}).
\]

Note that the operator $\hat{q}_{e_1, \ldots, e_D}$ is built from $D$ right-invariant vector fields. Since these are antisymmetric, $\hat{q}_{e_1, \ldots, e_D} = (-1)^D \hat{q}_{e_D, \ldots, e_1}$. In the case at hand, we have to use the projectors $\mathcal{P}_S$ to project on the allowed one-dimensional intertwiner space, the operator $\mathcal{P}_S \hat{q}^2 \mathcal{P}_S$ therefore has to vanish for the case $D + 1$ even (an antisymmetric matrix on a one-dimensional space is equal to 0). However, the volume operator depends on $\hat{q}^2$, and $\mathcal{P}_S \hat{q}^2 \mathcal{P}_S$ actually is a non-zero operator in any dimension, though trivially diagonal. Therefore, also $\hat{V}$ is diagonal.

The simplest non-trivial calculation involves a $D$-valent non-degenerate (i.e. no three tangents to edges at $v$ lie in the same plane) vertex $v$ where all edges are labelled by the defining representation of $SO(D + 1)$, and thus, the unique intertwiner which we will denote by $|N^{\delta(1)} \cdots N^{\delta(D)}\rangle$. We find

\[
\hat{q}_{e_1, \ldots, e_D}|N^{\delta(1)} \cdots N^{\delta(D)}\rangle = s(e_1, \ldots, e_D)\left(-\frac{1}{2}\right)^D |N_{\epsilon}^{\delta(1)} \cdots N^{\delta(D)}\rangle,
\]

\[
\hat{V}_v|N^{\delta(1)} \cdots N^{\delta(D)}\rangle = \left(\frac{1}{2}\right)^D D! \prod_{\delta(1) \cdots \delta(D)} |N^{\delta(1)} \cdots N^{\delta(D)}\rangle,
\]

i.e. for those special vertices, the volume operator preserves the simple vertices. For vertices of higher valence and/or other representations, we need to use the projectors. Of special interest are the vertices of valence $D + 1$ (triangulation) and 2D, where every edge has exactly one partner which is its analytic continuation through $v$ (cubulation). We find

\[
\hat{V}_v|N^{\delta(1)} \cdots N^{\delta(D+1)}\rangle = \left(\frac{1}{2}\right)^D (D + 1)! \prod_{\delta(1) \cdots \delta(D+1)} |N^{\delta(1)} \cdots N^{\delta(D+1)}\rangle,
\]

\[
\hat{V}_v|N^{\delta(1)} \cdots N^{\delta(D)}\rangle, \text{ cubic} = \left(\frac{1}{2}\right)^D (D)! \prod_{\delta(1) \cdots \delta(D)} |N^{\delta(1)} \cdots N^{\delta(D)}\rangle, \text{ cubic}.
\]

The dimensionality of the spatial slice now appears as a quantum number like the spins labelling the representations on the edges and it could be interesting to consider a large dimension limit in the spirit of the large $N$ limit in QCD.

(5) When introducing an Immirzi parameter in $D = 3$ [2], i.e. using the linear constraint $\epsilon_{JKL}^{\pi} \pi^{aKL} \approx 0$ while having $[h_{aIJ}(x), \pi^{bKL}(y)] = 2\pi^{aIJ} \delta_{bKL} \delta^{(D)}(x - y)$ with $\pi^{aIJ} = \pi^{aIJ} + 1/(2\gamma) \epsilon^{JKL} \pi^{aKL}$, the linear simplicity constraint operators become anomalous unless $\gamma = \pm \sqrt{\pi}$, the (anti-)self-dual case, which however results in non-invertibility of the prescription $\pi^{aIJ}$. Repeating the steps in section 3.1, we find that

\[\text{Up to } (N)^{D+1}, \text{ but in the chosen representation } \hat{N} \text{ acts by multiplication and therefore is less problematic than additional powers of right-invariant vector fields.}\]
these anomalous constraints require $\epsilon_{IJKL}N^I(R^{KL}_{\nu_1} - 1/(2\gamma)e^{KLMN}R^{MN}_{\nu_1}) \cdot f_\gamma = 0$. Since $\epsilon_{IJKL}N^I(R^{KL}_{\nu_1} - 1/(2\gamma)e^{KLMN}R^{MN}_{\nu_1})$ do not generate a subgroup, the constraint cannot be satisfied strongly if the edge $e$ transforms in an irreducible representation of $SO(D + 1)$ (by definition, the representation space does not contain an invariant vector).

In order to figure out the ‘correct’ quantization, one can try, in analogy to the strategy for the quadratic simplicity constraints, to weaken the imposition of the constraints at the quantum level. The basic difference between the linear and the quadratic simplicity constraints is that the time normal $N^I$ is left arbitrary in the quadratic case and fixed in the linear case. In order to lose this dependence in the linear case, one could average over all $N^I$ at each point in $\sigma$, which however leads to the Barrett–Crane intertwiners as described above. In analogy to the quadratic constraints, we could choose the subset

$$\epsilon_{IJKL}N^I(R^{KL}_{\nu_1} + R^{KL}_{\nu_2} + \cdots + R^{KL}_{\nu_{N-2}}) = 0$$

(3.19)

for each $N$-valent vertex plus the edge constraints. As above, the choice of the subset specifies a recoupling scheme and the imposition of the constraints leads to the contraction of the virtual edges and virtual intertwiners of the recoupling scheme with the $SO(D)_N$-invariant vectors $\xi_{\nu_1}(N)$ and their complex conjugates $\bar{\xi}_{\nu_1}(N)$, see figure 2. Gauge invariance can still be used at each (virtual) vertex in this calculation in the form $\sum \bar{R}_e = 0$, which is sufficient since only $\bar{R}_e$ appears in the linear simplicity constraints. If we now integrate over each pair of $\bar{\xi}_{\nu_1}(N)$ ‘generated’ by the elements of the proposed subset of the simplicity constraint operators separately, we obtain projectors on simple representations for each of the virtual edges in the recoupling scheme. The integration over $N^I$ for the edge constraints yields projectors on simple representations in the same manner. Finally, we obtain the simple intertwiners of the quadratic operators in addition to solutions where incoming edges are contracted with $SO(D)_N$-invariant vectors $\xi_{\nu_1}(N)$. A few remarks are appropriate.

(6) Although this procedure yields a promising result, it contains several non-standard and ahoc steps which have to be justified. One could argue that the ‘correct’ quantization of the linear and quadratic simplicity constraints should give the same quantum theory; however,
as is well known, classically equivalent theories result in general in non-equivalent quantum theories, which nevertheless can have the same classical limit.

(7) It is unclear how to proceed with ‘integrating out’ \( N^I \) in the general case. For the vacuum theory, integration over every point in \( \sigma \) gives the Barrett–Crane intertwiner for the edges contracted with \( \text{SO}(D)_\mathbb{N} \)-invariant vectors. This type of integration would also get rid of the one-valent vertices and thus allow for a natural unitary map to the quadratic solutions as already mentioned above.

(8) When introducing fermions, there is the possibility for non-trivial gauge-invariant functions of \( N^I \) at the vertices which immediately results in the question of how to integrate out this \( N^I \)-dependence. Next to including those \( N^I \) in the above integration or to integrate out the remaining \( N^I \) separately, one could transfer this integration back into the scalar product. Since the authors are presently not aware of an obvious way to decide about these issues, we will leave them for further research.

3.5. Mixed quantization

Since the implementation of the quadratic simplicity constraints described above yields a more promising result than the implementation of the linear constraints, we can try to perform a mixed quantization by noting that we can classically express the linear constraints for even \( D \) in the form

\[
\frac{1}{4} \epsilon_{IJKL} \pi^{al} \pi^{bk} \approx 0, \quad N^I - n^I(\pi) \approx 0.
\]

The phase space extension derived in [5] remains valid when interchanging the linear simplicity constraint for the above constraints. The reason for restricting \( D \) to be even is that we have an explicit expression for \( n^I(\pi) \), see [1, 2]. Since a quantization of \( n^I(\pi) \) will most likely not commute with the Hamiltonian constraint operator, we resort to a master constraint. Note that the expression

\[
M'_N := \left( (N^I - n^I(\pi)) \sqrt{q}^{-D-1} \delta_{IJ} ((N^J - n^J(\pi)) \sqrt{q}^{-D-1}) \right) \sqrt{q}^{-1},
\]

which is the densitized square of \( N^I - n^I(\pi) \), can be quantized as

\[
\hat{M}'_N = 2 \sqrt{q}^{-2D} \left( \sqrt{|\hat{V}^I\hat{V}_I|} - N_i \hat{V}^I \right),
\]

when using a suitable factor ordering, where the quantization of \( \sqrt{q}^{-2D} \) is described in [3]. The solution space is not empty since the intertwiner

\[
s(e_1, ..., e_D|N^A_1, N^A_2, ..., N^A_D) + |N_\pi e^{B_1, ..., A_D}\rangle
\]

is annihilated by \( \hat{M}_N \), which can be easily checked when using the results of the volume operator acting on the solution space of the full set of linear simplicity constraint operators. In order to turn the expression into a well-defined master constraint operator, we have to square it again and to adjust the density weight, leading to

\[
\hat{M}_N = 4 \left( \sqrt{q}^{5/2} \left( \sqrt{|\hat{V}^I\hat{V}_I|} - N_i \hat{V}^I \right) \right) \sqrt{q}^{5/2} \left( \sqrt{|\hat{V}^I\hat{V}_I|} - N_i \hat{V}^I \right),
\]

which is by construction a self-adjoint operator with non-negative spectrum. We remark that it was necessary to use the fourth power of the classical constraint for quantization, because the second power, having the desired property that its solution space is not empty, does not qualify as a well-defined master constraint operator in the ordering we have chosen. There exists however no a priori reason why one should not take into account master constraint operators constructed from higher powers of classical constraints [49]. Curiously, the quadratic
simplicity constraint operators as given above do not annihilate the solution displayed. Clearly, the calculations will become much harder as soon as vertices with a valence higher than $D$ are used, since the building blocks of the volume operator will not be diagonal on the intertwiner space.

This type of quantization is further discussed in section 4.3, where a possible application to using EPRL intertwiners is outlined. In contrast to the earlier assumption of $D$ being even in order to have an explicit expression $n^I(\pi)$, we can also perform the mixed quantization using $n^I n_J \approx \frac{1}{2} \left( \pi^{aIJ} \pi_{aIKL} - \zeta_{IJ} \right)$ for the Euclidean internal signature $\zeta = 1$ and the constraint $N_J (n^I n^I - N^I N^I) \approx 0$. For the application proposed, we will only need that the corresponding master constraint can be regularized such that it vanishes when not acting on non-trivial vertices, which can be achieved as before.

4. Comparison to existing approaches

In this section, we are going to comment on the relation of existing results from the spin foam literature to the proposals in this paper. In short, the main conclusion will be that in the case of four spacetime dimensions, many results from the spin foam literature can also be used in the canonical framework. However, they fail to work in higher dimensions due to special properties of the four-dimensional rotation group which are heavily used in spin foams. We will not comment on results based on coherent state techniques [16, 33–35] since we do not see a resemblance to our results which do not make use of coherent states in any way. Nevertheless, similarities could be present as the relation between the EPRL [24] and the FK [16] models show.

4.1. Continuum versus discrete starting point

The starting point for introducing the simplicity constraints in the spin foam models is the reformulation of general relativity as a BF theory subject to the simplicity constraints, and thus similar to the point of view taken in this series of papers. The crucial difference however is that while spin foam models start classically from discretized general relativity, the canonical approach discussed here starts from its continuum formulation. When looking at the simplicity constraints, this difference manifests itself in the choice of $(D-1)$-surfaces over which the generalized vielbeins (i.e. the bivectors in spin foam models) have to be smeared. Starting from a discretization of spacetime, the set of $(D-1)$-surfaces is fixed by foliating the discretized spacetime. Restricting to a simplicial decomposition of a four-dimensional spacetime as an example, these would e.g. be the faces $f$ of a tetrahedron $t$ in the boundary of the discretization. It follows that one can take the bivectors $B^I_J$ integrated over the individual faces of a tetrahedron, $B^I_J (t) := \int_f B^I_J$, as the basic variables and the quadratic (off)-diagonal simplicity constraints read [24]

$$C_{ff} := \epsilon_{IJKL} B_f (t)^{IJ} B_f (t)^{KL} \quad \text{diagonal simplicity } \forall f \in t,$$  

$$C_{ff'} := \epsilon_{IJKL} B_f (t)^{IJ} B_{f'} (t)^{KL} \quad \text{off-diagonal simplicity } \forall f, f' \in t.$$  

In the continuum formulation, however, we have to consider all possible $(D-1)$-surfaces, and thus also hypersurfaces containing the vertex $v$ dual to the tetrahedron $t$. The resulting flux operators a priori contain a sum of right-invariant vector fields $R^I_e$ acting on all the edges $e$ connected to $v$. While this poses no problem for the diagonal simplicity constraints which act on edges of the spin networks as shown in [3], the off-diagonal simplicity constraints arising when both surfaces contain $v$ are not given by (4.2), but by sums over different $C_{ff'}$, see.
[3] for details. It can however be shown by suitable superpositions of simplicity constraints associated with different surfaces that (4.2) is actually implied also by the quadratic simplicity constraints arising from a proper regularization in the canonical framework. This statement is non-trivial and had to be proved in [3]. Thus, we can also in the canonical theory consider the individual building blocks (2.2) as in section 2 of this paper. Furthermore, the same is also true when using linear simplicity constraints, i.e. the properly regularized linear simplicity constraints in the canonical theory imply that all building blocks (3.3) vanish.

We also note that there is no analogue of the normalization simplicity constraints [11] in the canonical treatment since the generalized vielbeins do not have time-like tensorial indices after being pulled back to the spatial hypersurfaces.

4.2. Projected spin networks

Projected spin networks were originally introduced in [50, 51] to describe Lorentz covariant formulations of loop quantum gravity, meaning that the internal gauge group is $SO(1, 3)$ (or $SL(2, \mathbb{C})$) instead of $SU(2)$. The basic idea is that next to the connection, the time normal field, often called $x$ or $\chi$ in the spin foam literature, becomes a multiplication operator since it Poisson commutes classically with the connection. Since the physical degrees of freedom of loop quantum gravity formulated in terms of the usual $SU(2)$ connection and its conjugate momentum are orthogonal to the time normal field, one performs projections in the spin networks from the full gauge group $SO(1, 3)$ to a subgroup stabilizing the time normal. Since the projector transforms covariantly under $SO(1, 3)$, a (gauge-invariant) projected spin network is already defined by its evaluation for a specific choice of the time normals and the resulting effective gauge invariance is only $SU(2)$, which exemplifies the relation to the usual $SU(2)$ formulation in the time gauge $x^I = (NI) = (1, 0, 0, 0)$.

Despite its close relation to the techniques used in this paper and its merits for the four-dimensional treatment, there are several problems connected with using this approach in the canonical framework discussed in this series of papers which we will explain now. While the extension of projected spin networks to different gauge groups has already been discussed in [51], there is a subtle problem associated with the part of the connection which is projected out by the projections, which could not have been anticipated by looking at loop quantum gravity in terms of the Ashtekar–Barbero variables. There, the physical information in the connection, the extrinsic curvature, is located in the rotational components of the connection. To see this, consider in four dimensions the two-parameter family of connections discussed in [2],

\[ A_{aIJ} = A^{\text{hyb}}_{aIJ} + \beta K_{aIJ} + \gamma \frac{1}{2} \epsilon_{ijkl} K_{aKL}, \]  

\[ (4.3) \]

where $\gamma$ corresponds to the Barbero–Immirzi parameter restricted to four dimensions and $\beta$ is the new free parameter appearing in any dimension. $K_{aIJ}$ decomposes as [1, 2]

\[ K_{aIJ} = 2N^j K_{aIJ} + K_{aIJ}^{\text{trace}} + K_{aIJ}^{\text{trace free}}, \]  

\[ (4.4) \]

where $K_{aIJ}$ means that $N^k \tilde{K}_{aIJ} = N^j \tilde{K}_{aIJ} = 0$ and the trace/traceless split is performed with respect to the hybrid vielbein. The extrinsic curvature which we need to recover from $A_{aIJ}$ is located in $\tilde{K}_{aIJ}$, whereas $K_{aIJ}^{\text{trace}}$ vanishes by the Gauß constraint and $K_{aIJ}^{\text{trace free}}$ is pure gauge from the simplicity gauge transformations.

Now setting $\beta = 0$ and $N^d = (1, 0, 0, 0)$ in four dimensions, we recover the Ashtekar–Barbero connection and see that the physical information is located in the rotational components of $A_{aIJ}$. It thus makes sense to project onto this subspace in the projected spin network formulation.

\[ A_{aIJ}^{\text{projected}} = \frac{1}{2} \epsilon_{ijkl} K_{aKL} \]  

\[ (4.5) \]

Note that the definitions of the parameters are different in [2] for calculational simplicity, but here we prefer this parametrization to make our point clear.
network construction, i.e. we are not losing physical information. On the other hand, setting $\gamma = 0$ in four dimensions or going to higher dimensions, we see that a projection onto the subspace orthogonal to $N^l$ annihilates the physical components of the connection. This would not be necessarily an issue if one would just project the projected spin network at the intertwiners, but when one tries to go to fully projected spin networks as proposed in [50]. Then, since one would take a limit of inserting projectors at every point of the spin network, the physical information in the connection would be completely lost.

Next to this problem, there are other problems associated with taking an infinite refinement limit for projected spin networks as discussed by Alexandrov [50] and Livine [51], e.g., that fully projected spin networks are no longer spin networks (since they only contain vertices and no edges) and, connected with this problem, that the trivial bivalent vertex, the Kronecker delta, is not an allowed intertwiner. Similar problems have been encountered in section 3, i.e. while the vertex simplicity constraints could be solved by a construction very similar to projected spin networks where one projects the incoming and outgoing edges at the intertwiner in the direction of the time normal $N^l$, imposing the linear simplicity constraint on the edges, one would have to insert ‘trivial’ bivalent vertices of the form $N^l N^j$ at every point of the spin network, whereas one would need to insert the Kronecker delta $\delta^l l$ to achieve cylindrical consistency while maintaining a spin network containing edges and not only vertices.

Thus, the main problem with using (fully) projected spin networks is connected to the fact that we do not know of an analogue of the Barbero–Immirzi parameter in higher dimensions which would allow us to put the extrinsic curvature also in the rotational components of the connection. In four dimensions on the other hand, this problem would be absent and one would be left with the issue of refining the projected spin networks, which is however also present in section 3 of this paper. Therefore, using projected spin networks in four dimensions with a non-vanishing Barbero–Immirzi parameter is an option for the canonical framework developed in this series of papers and the known issues discussed above should be addressed in further research.

4.3. EPRL model

The basic idea of the EPRL model is to implement the diagonal simplicity constraints as usual, but to replace the off-diagonal simplicity constraints by linear simplicity constraints which are implemented with a master constraint construction [24] or weakly [52]. Furthermore, the Barbero–Immirzi parameter is a necessary ingredient. We restrict here to the Euclidean model since its group theory is much closer to the connection formulation with compact gauge group $SO(D + 1)$. While the diagonal simplicity constraints give the well-known relation

$$ (j^+)^2 = \left(\frac{\gamma + 1}{\gamma - 1}\right)^2 (j^-)^2, $$

the master constraint for the linear constraints gives [24], up to $\hbar$ corrections\(^{11}\),

$$ k^2 = \left(\frac{2j^-}{1 - \gamma}\right)^2 = \left(\frac{2j^+}{1 + \gamma}\right)^2, $$

where $k$ is the quantum number associated with the Casimir operator of the $SU(2)$ subgroup stabilizing $N^l$. Depending on the value of the Barbero–Immirzi parameter, either $k = j^+ + j^-$ or $k = |j^+ - j^-|$ is selected by this constraint. The EPRL intertwiner for $SO(4)$ spin networks

\(^{11}\)Note that these $\hbar$ corrections are necessary since the master constraint, by construction, has the same solution space as the original constraint [49], i.e. $\hat{C} \hat{\psi} = 0$ implies $\hat{C} |\psi\rangle = 0$. In the master constraint language, one subtracts an operator from the master constraint which vanishes in the classical limit to obtain a sufficiently large solution space.
with arbitrary valency [29] is then constructed by first coupling the two SU(2) subgroups of SO(4) holonomies in the representations \((j^+, j^-)\), calculated along incoming and outgoing edges to the intertwiner, to the \(k\) representation. Then, the \(k\) representations associated with each edge are coupled via an SU(2) intertwiner and the complete construction is then projected into the set of SO(4) intertwiners.

An alternative derivation proposed by Ding and Rovelli [52] makes use of weakly implementing the linear simplicity constraints, i.e. restricting to a subspace \(\mathcal{H}^{\text{ext}}\) such that

\[
\langle \phi | \hat{C} | \psi \rangle = 0 \quad \forall |\phi\rangle, |\psi\rangle \in \mathcal{H}^{\text{ext}}.
\] (4.7)

In this approach, one can also show that the volume operator restricted to \(\mathcal{H}^{\text{ext}}\) has the same spectrum as in the canonical theory, which is an important test to establish a relation between the canonical theory and the EPRL model.

Closely related to what we already observed in the previous subsection on projected spin networks, the EPRL model makes heavy use of the fact that SO(4) splits into two SU(2) subgroups and that the Barbero–Immirzi parameter is available in four dimensions. Thus, we would have to restrict to four dimensions with non-vanishing \(\gamma\) if we would want to use the EPRL solution to the simplicity constraints. One upside of this solution when comparing to our proposition for solving the quadratic constraints is that no choice problem occurs, i.e. if we map the quantum numbers of the EPRL intertwiners to SU(2) spin networks, a change of recoupling basis in the SU(2) spin networks results again in EPRL intertwiners solving the same simplicity constraints. The problem of stability of the solution space \(\mathcal{H}^{\text{ext}}\) of the simplicity constraint under the action of the Hamiltonian constraint is however, to the best of our knowledge, not circumvented when using EPRL intertwiners.

Also, in order to use the EPRL solution in the canonical framework, one would have to discuss exactly what it means to use linear and quadratic simplicity constraints in the same formulation, i.e. if one can freely interchange them and how continuity of the time normal field is guaranteed at the classical level if one changes from the quadratic constraints to linear constraints from one point on the spatial hypersurface to another. The mixed quantization proposed in section 3.5 can be seen as an attempt to using both the time normal as an independent variable and quadratic simplicity constraints. In this case, the main difference is the presence of an additional constraint relating the time normal constructed from the generalized vielbeins to the independent time normal (which could be used in the linear simplicity constraints). In section 3.5, this additional constraint was regularized as a master constraint which acts only on vertices. Taking the point of view that one can freely change between using the quadratic constraints plus this additional constraint or the linear constraints, one could choose the linear constraints for vertices and the quadratic constraints for edges. Since we can use a factor ordering for the master constraint where a commutator between a holonomy and a volume operator is ordered to the right, the master constraint would vanish on edges and only the quadratic simplicity constraints would have to be implemented, which are however not problematic. At vertices, we would be left with the linear constraints and could use the EPRL intertwiners. Thus, the EPRL solution seems to be a viable option in four dimensions. Whether one considers it natural or not to use both linear and quadratic constraints in the same formulation is a matter of personal taste. Nevertheless, it would be desirable to have only one kind of simplicity constraints.

A further comment is due on the starting point of spin foam models, which is a BF-theory subject to simplicity constraints. It has been argued by Alexandrov [53] that the secondary constraints resulting from the canonical analysis, i.e. the \(D_M^a\) -constraints on the torsion of \(A_{aIJ}\) from our companion paper [2], should be taken into account also in spin foam models.
In the present canonical formulation, these constraints were removed by the gauge unfixing procedure [2] and thus do not have to be taken into account here. The requirement for the validity of this step was to modify the Hamiltonian constraint by an additional term quadratic in the $D_{\alpha}^{\mu}$-constraints (the gauge unfixing term) which renders the simplicity constraints stable. While this ensures that we have to deal only with the non-commutativity of the (singularly smeared, or quantum) simplicity constraints in this paper, the converse does not necessarily follow. Since the Hamiltonian constraint that one obtains from the canonical analysis of BF-theory subject to simplicity constraints, the classical starting point of spin foam models, is not the modified Hamiltonian constraint considered here, but the one which results in the secondary $D_{\alpha}^{\mu}$-constraints, it does not follow that these secondary constraints do not have to be taken into account in spin foam models. On the other hand, the present formulation hints that it might be possible to construct a spin foam model subject to simplicity constraints (and not $D_{\alpha}^{\mu}$-constraints) which coincides with the dynamics defined by the modified Hamiltonian constraint. In fact, it was recently shown that the transfer operator of spin foam models can be written as $T = \mu^\dagger W \mu$ (here for the EPRL model) [54], where $\mu$ projects onto the solution space of the simplicity constraint. Taking into account the philosophy of spin foam models to impose the simplicity constraint at every time step in order to ensure that the second class $D_{\alpha}^{\mu}$-constraints are satisfied, it is conceivable that the gauge unfixing term of the Hamiltonian constraint in [1, 2] could emerge from these $\mu$-projections when taking the continuum limit of the spin foam transfer operator. Thus, in the light of plausible arguments for both sides, only an explicit calculation will be able to decide this issue.

As a last remark, we point out that the non-commutativity of the linear simplicity constraints in the EPRL model results from using $\gamma \neq 0$, and thus, we are not faced with this problem in higher dimensions. Essentially, as discussed in more detailed in remark 3.4 of section 3.4, while the rotations stabilizing $N^i$ form an $SO(D)$ subgroup of $SO(D + 1)$, the linear simplicity constraints in four dimensions with $\gamma \neq 0$ and $\beta \neq 0$ do not generate such a subgroup.

5. Discussion and conclusions

Let us briefly discuss the results of this paper and judge the different approaches.

First, the mechanism for avoiding the non-commutativity in the quadratic simplicity constraints discussed in section 2 is new to the best of our knowledge and we do not see any indication that the solution space is identical to previous results (up to the fact that it has the same ‘size’ as SU(2) spin networks). In the spin foam literature, the linear simplicity constraints are the cornerstones of the new spin foam models and have been introduced since the quadratic simplicity constraints acting on vertices do not commute. While the methods for treating supergravity discussed in [5] necessarily need an independent time normal and thus suggest using linear simplicity constraints, there is no need for the linear constraints in pure gravity (except for the fact that they exclude the topological sector in four dimensions). Therefore, one should not dismiss the quadratic constraints, especially since the linear constraints come with their own problems in the canonical approach. The solution presented in section 2 is certainly not free of problems, most prominently the choice of the maximal commuting subset, but its close relation the SU(2)-based theory and the (natural) unitarity of the intertwiner map to SU(2) intertwiners make it look very promising.

The linear simplicity constraints come with their own set of problems, many of which were already known in the spin foam literature. While the results of section 2 would naturally lead us to consider the quadratic constraints, the connection formulation of higher dimensional supergravity developed in [5] makes it necessary to use an independent time normal as an
additional phase space variable. This time normal would naturally point towards using linear simplicity constraints, although the mixed quantization of section 3.5 could avoid this. Since there is no anomaly appearing when using the linear simplicity constraints (with $\gamma = 0$ in four dimensions), we should implement them strongly. However, this leads to a solution space very different from the SU(2) spin networks. At this point, it seems to be the best to let oneself be guided by physical intuition and the results from the quadratic simplicity constraints as well as the desired resemblance to SU(2) spin networks. Ad hoc methods for getting close to this goal have been discussed in section 3.4. We however stress that these methods are, as said, ad hoc and they do not follow from standard quantization procedures. The mixed quantization discussed at the end of section 3 also does not seem completely satisfactory, especially since the master constraint ensuring the equality of the independent normal $N^I$ and the derived normal $n^I(\pi)$ is very complicated to solve. Nevertheless, in section 4.3, an application to EPRL intertwiners is outlined which could avoid this problem by using linear simplicity constraints for the vertices. The strength of the mixed quantization is thus that it provides a mechanism to incorporate both the quadratic simplicity constraints and an independent time normal in the same canonical framework, which is what performed on the path integral side in the EPRL model. A comparison to results from the spin foam literature, especially projected spin networks and the EPRL model, shows that many of the problems connected with using the linear simplicity constraints have already been known, partly in different guises. While using these known results in our framework seems to be a viable option in four dimensions, we are unaware of possible ways to extend them also to higher dimensions since main ingredients are a non-vanishing Barbero–Immirzi parameter as well as special properties of SO(4).

In conclusion, we reported on several new ideas of how to treat the simplicity constraints which appear in our connection formulation of general relativity in any dimension $D \geq 3$ [1–6] and found that none of the presented ideas are entirely satisfactory at this point and further research on the open questions needs to be conducted. We hope that the discussion presented in this paper will be useful for an eventually consistent formulation.

Acknowledgments

NB and AT thank Emanuele Alesci, Jonathan Engle, Alexander Stottmeister and Antonia Zipfel for numerous discussions and Karl-Hermann Neeb and Toshiyuki Kobayashi for the counsel on representation theory. NB and AT also thank the German National Merit Foundation for financial support. The part of the research performed at the Perimeter Institute for Theoretical Physics was supported in part by funds from the Government of Canada through NSERC and from the Province of Ontario through MEDT. Final improvements of this work were supported by the NSF grant PHY-1205388 and the Eberly research funds of The Pennsylvania State University. We acknowledge helpful suggestions from the referees of this paper which greatly improved its readability.

Appendix. Simple irreps of $\text{SO}(D+1)$ and square integrable functions on the sphere $S^D$

There is a natural action of $\text{SO}(D+1)$ on $F \in \mathcal{H} := L_2(S^D, d\mu)$ given by $\pi(g)F(N) := F(g^{-1}N)$. The $\pi(g)$ are called quasi-regular representations of $\text{SO}(D+1)$. The generators in this representation are of the form $\tau_{ij} = \frac{1}{4}(\frac{1}{\sqrt{N}}N_j - \frac{1}{\sqrt{N}}N_i)$ and are known to satisfy the quadratic simplicity constraint $\tau_{ijkl}\tau_{kl} = 0$ [11]. These representations are reducible. The
representation space can be decomposed into spaces of harmonic homogeneous polynomials \( \mathcal{Y}^{D+1,l} \) of degree \( l \) in \( D + 1 \) variables, \( \mathcal{L}_2(\mathcal{Y}^{D}) = \sum_{l=0}^{\infty} \mathcal{Y}^{D+1,l} \). The restriction of \( \pi(g) \) to these subspaces gives irreducible representations of SO\( (D + 1) \) with highest weights \( \Lambda = (l, 0, \ldots, 0), l \in \mathbb{N} \). These are (up to equivalence) the only irreducible representations of SO\( (D + 1) \) satisfying the quadratic simplicity constraint [11] and therefore are mostly called simple representations in the spin foam community, which we adopted in this work. Note that these representations have been studied quite extensively in the mathematical literature, where they are called most degenerate representations [55–57], (completely) symmetric representations [56, 58–60] or representations of class one (with respect to an SO\( (D) \) subgroup) [43]. The latter is due to the fact that these representations of SO\( (D + 1) \) are the only ones which have in their representations space a non-zero vector invariant under an SO\( (D) \) subgroup, which is exactly the definition of being of class one w.r.t. a subgroup given in [43]. An orthonormal basis in \( \mathcal{Y}^{D+1,l} \) is given by generalizations of spherical harmonics to higher dimensions [43] which we denote \( \mathcal{Z}_l^\mathcal{K}(N) \),

\[
\int_{\mathcal{Y}^{D}} \mathcal{Z}_l^{\mathcal{K}}(N) \mathcal{Z}_l^{\mathcal{M}}(N) \, dN = \delta_l^{\mathcal{K}} \delta_l^{\mathcal{M}},
\]

(A.1)

where \( \mathcal{K} \) denotes an integer sequence \( \mathcal{K} := (K_1, \ldots, K_{D-2}, \pm K_{D-1}) \) satisfying \( l \geq K_1 \geq \cdots \geq K_{D-1} \geq 0 \) and analogously defined \( M \). \( F_l(N) \in \mathcal{Y}^{D+1,l} \) can be decomposed as \( F_l(N) = \sum_k a_k^{\mathcal{K}} \mathcal{Z}_l^{\mathcal{K}}(N) \) where the sum runs over those integer sequences \( \mathcal{K} \) allowed by the above inequality. Since \( \mathcal{L}_2(\mathcal{Y}^{D}) = \sum_{l=0}^{\infty} \mathcal{Y}^{D+1,l} \), any square integrable function \( F(N) \) on the sphere can be expanded in a mean-convergent series of the form [43]

\[
F(N) = \sum_{l=0}^{\infty} \sum_{K_l} a_{K_l}^{\mathcal{K}} \mathcal{Z}_l^{\mathcal{K}}(N).
\]

(A.2)

Consider a recoupling basis [61] for the ONB of the tensor product of \( N \) irreps. Choose a labelling of the irreps \( \tilde{\Lambda}_1, \ldots, \tilde{\Lambda}_N \). Then, consider the ONB

\[
|\tilde{\Lambda}_1, \ldots, \tilde{\Lambda}_N; \tilde{K}_1, \tilde{K}_2, \ldots, \tilde{K}_{D-1}; \tilde{K}, \tilde{M}\rangle,
\]

(A.3)

(with certain restrictions on the values of the intermediate and final highest weights). A basis in the intertwiner space is given by

\[
|\tilde{\Lambda}_1, \ldots, \tilde{\Lambda}_N; \tilde{K}_1, \tilde{K}_2, \ldots, \tilde{K}_{D-1}; 0, 0\rangle
\]

(A.4)

(with certain restrictions). A change of recoupling scheme corresponds to a change of basis in the intertwiner space. A basis in the intertwiner space of \( N \) simple irreps is given by

\[
|\Lambda_1, \ldots, \Lambda_N; \Lambda_{12}, \Lambda_{13}, \ldots, \Lambda_{1N-1}; 0, 0\rangle
\]

(A.5)

(with certain restrictions), since in the tensor product of two simple irreps, non-simple irreps appear in general [60, 59],

\[
\langle \lambda_1, 0, \ldots, 0 \rangle \otimes \langle \lambda_2, 0, \ldots, 0 \rangle = \sum_{k=0}^{\lambda_2} \sum_{l=0}^{\lambda_2-k} \langle \lambda_1 + \lambda_2 - 2k - l, l, 0, \ldots, 0 \rangle \quad (\lambda_2 \leq \lambda_1).
\]

(A.6)

References

[1] Bodendorfer N, Thiemann T and Thurn A 2013 New variables for classical and quantum gravity in all dimensions: part I. Hamiltonian analysis Class. Quantum Grav. 30 045001 (arXiv:1105.3703 [gr-qc])

[2] Bodendorfer N, Thiemann T and Thurn A 2013 New variables for classical and quantum gravity in all dimensions: part II. Lagrangian analysis Class. Quantum Grav. 30 045002 (arXiv:1105.3704 [gr-qc])
[34] Dupuis M and Livine E R 2011 Holomorphic simplicity constraints for 4D spinfoam models Class. Quantum Grav. 28 215022 (arXiv:1104.3683 [gr-qc])
[35] Dupuis M, Freidel L, Livine E R and Speziale S 2012 Holomorphic Lorentzian simplicity constraints J. Math. Phys. 53 032502 (arXiv:1107.5274 [gr-qc])
[36] Mitra P and Rajaraman R 1989 Gauge-invariant reformulation of an anomalous gauge theory Phys. Lett. B 225 267–71
[37] Jackiw R and Rajaraman R 1985 Vector-meson mass generation by chiral anomalies Phys. Rev. Lett. 54 1219–21
[38] Lott J and Rajaraman R 1985 Degrees of freedom and the quantization of anomalous gauge theories Phys. Lett. B 165 321–6
[39] Faddeev L 1984 Operator anomaly for the gauss law Phys. Lett. B 145 81–4
[40] Rajaraman R 1985 Hamiltonian formulation of the anomalous chiral Schwinger model Phys. Lett. B 154 305–9
[41] Flori C and Thiemann T 2008 Semiclassical analysis of the loop quantum gravity volume operator: part I. Flux coherent states arXiv:0812.1537 [gr-qc]
[42] Rovelli C 2011 Zakopane lectures on loop gravity arXiv: 1102.3660 [gr-qc]
[43] Vilenkin N J 1968 Special Functions and the Theory of Group Representations (Providence, RI: American Mathematical Society)
[44] Thiemann T 1998 Kinematical Hilbert spaces for Fermionic and Higgs quantum field theories Class. Quantum Grav. 15 1487–512 (arXiv:gr-qc/9705021)
[45] Higuchi A 1991 Quantum linearization instabilities of de Sitter spacetime: part II Class. Quantum Grav. 8 1983–2004
[46] Giulini D and Marolf D 1999 On the generality of refined algebraic quantization Class. Quantum Grav. 16 2479–88 (arXiv:gr-qc/9812024)
[47] Giulini D and Marolf D 1999 A uniqueness theorem for constraint quantization Class. Quantum Grav. 16 2489–505 (arXiv:gr-qc/9902045)
[48] 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 (arXiv:1108.1178 [gr-qc])
[49] Thiemann T 2006 Quantum spin dynamics: part VIII. The master constraint Class. Quantum Grav. 23 2249–65 (arXiv:gr-qc/0510011)
[50] Alexandrov S 2002 Hilbert space structure of covariant loop quantum gravity Phys. Rev. D 66 024028 (arXiv:gr-qc/0201087)
[51] Livine E R 2002 Projected spin networks for Lorentz connection: linking spin foams and loop gravity Class. Quantum Grav. 19 5525–41 (arXiv:gr-qc/0207084)
[52] Ding Y and Rovelli C 2010 The volume operator in covariant quantum gravity Class. Quantum Grav. 27 165003 (arXiv:0911.0543 [gr-qc])
[53] Alexandrov S 2010 New vertices and canonical quantization Phys. Rev. D 82 024024 (arXiv:1004.2260 [gr-qc])
[54] Dittrich B, Hellmann F and Kaminski W 2012 Holonomy spin foam models: boundary Hilbert spaces and time evolution operators arXiv:1209.4539 [gr-qc]
[55] Hormeß M and Junker G 1999 More on coupling coefficients for the most degenerate representations of SO(n) J. Phys. A: Math. Gen. 32 4249–56 (arXiv:math-ph/9903035)
[56] Alisauskas S 2002 Coupling coefficients of SO(n) and integrals involving Jacobi and Gegenbauer polynomials J. Phys. A: Math. Gen. 35 7323–45 (arXiv:math-ph/0210148)
[57] Junker G 1993 Explicit evaluation of coupling coefficients for the most degenerate representations of SO(n) J. Phys. A: Math. Gen. 26 1649–61
[58] Alisauskas S 1987 Some coupling and recoupling coefficients for symmetric representations of SO(n) J. Phys. A: Math. Gen. 20 35
[59] Girardi G, Sciarrino A and Sorba P 1982 Generalized Young tableaux and Kronecker products of SO(n) representations Physica A 114 365–9
[60] Girardi G, Sciarrino A and Sorba P 1982 Kronecker products for SO(2p) representations J. Phys. A: Math. Gen. 15 1119–29
[61] Agrawala V and Belfiante J 1968 Graphical formulation of recoupling theory for any compact group Ann. Phys. 49 130–70