Gauge field theories are an important cornerstone of modern physics and encompass the fundamental forces of nature, including electromagnetism and nuclear forces. The physical information is captured in Wilson loops [1], or holonomies, which describe how a quantity is parallel transported along a given closed path. Local gauge transformations can modify the fundamental fields independently at each space-time point but leave any traced Wilson loop invariant. On the lattice, gauge invariant observables are typically formulated in terms of traced Wilson loops of different shapes. The most basic example is the Wilson action, which is formulated entirely in terms of 1×1 loops, so-called plaquettes. The Wilson action can be systematically improved by including terms involving larger loops [2–6]. Planar rectangular loops are used for characterizing confinement. Most famously, the potential of a static quark pair can be computed from the expectation value of a Wilson loop with large extent in the temporal direction [7]. Improved approximations to the energy momentum tensor or the topological charge density can involve also nonplanar loops of growing size [8–10]. As the number of possible loops on a lattice grows exponentially with its path length, a systematic treatment of higher order contributions can become increasingly challenging.

Artificial neural networks provide a way to automatically extract relevant information from large amounts of data. They have become increasingly popular in many Abelian lattice applications, such as for ϕ^4 scalar field, Ising, XY, Potts and Yukawa models, where they can recognize classical [11] and topological [12] phase transitions from field configurations, determine local and nonlocal features [13,14], or infer action parameters [15]. Neural networks can improve the efficiency of sampling techniques [16], extract optimal renormalization group transformations [17], or reconstruct spectral functions from Green's functions [18]. By the universal approximation theorem, these networks can, in principle, learn any function [19–21]. In order to avoid merely memorizing training samples, imposing additional restrictions on these networks can improve their generalization capabilities [22]. Global translational equivariance induces convolutions [23], which form the basis of convolutional neural networks (CNNs). Additional global symmetry groups, such as global rotations, can be incorporated using group equivariant CNNs (G-CNNs) [24–29]. This approach can be extended to local gauge symmetries. Even though gauge invariant observables can be learned to some extent by nonequivariant networks [30], recently there has been a lot of interest in incorporating gauge symmetries directly into the network structure. For discrete ones, equivariant network structures have been implemented for the icosahedral group [31] or for the Z_2 gauge group [32]; for continuous ones, a much larger symmetry space is available [33]. A recent seminal work demonstrated that incorporating U(1) or $SU(N_c)$ gauge symmetries into a neural network can render flow-based sampling orders of magnitude faster than traditional approaches [34,35]. This impressive result was obtained using parametrized invertible coupling layers that essentially depend on parallel-transported plaquettes. Up until now, machine learning applications that require larger Wilson loops have relied on manually picking a set of relevant Wilson loops [36] or on simplifications due to the choice of a discrete Abelian gauge group [37]. A comprehensive treatment for continuous non-Abelian gauge groups has been missing so far, and there is an obvious desire to systematically generate all Wilson loops from simple local operations.