Ansatz	Surjective	Lo
$\mathbb{SU}(d)$-gate [?, ?, ?, ?, ?]	yes	
Overparameterized $\mathbb{SU}(d)$-gate	yes	
product-of-exponentials [?, ?, ?, ?, ?, ?, ?, ?, ?, ?]	no*	
Overparameterized product of exponentials [?, ?, ?, ?, ?, ?]	yes, with enough parameters	
Composite (Eq. (??))	yes	
Cayley transform	yes**	

Table 1: Comparison of the different ansätze described in this work and their properties. Surjectivity means that any unitary in SU(d) can be reached, while local surjectivity means that the partial derivatives span the tangent space of SU(d) at any point in the parameter landscape (cf. def:local$_s urjectivity).The$$\mathbb{SU}(d)$-gate ansatz requires d^2-1 parameters, is surjective on SU(d) but is never locally surjective even with the use of overparameterization. The product-of-exponentials ansatz is in general (*) not surjective for all dimensions d, is not locally surjective and typically uses less than d^2-1 parameters. With sufficient overparameterization it becomes surjective [?], however it remains unknown if it also becomes locally surjective. The composite ansätze introduced in this work are both surjective and locally surjective and require $2(d^2-1)$ parameters. The Cayley transform is technically speaking not surjective onto SU(d), (**) but can still reach any unitary up to an irrelevant global phase. It is locally surjective and uses d^2 parameters. Hence, the requirements for the main theorem hold only for the composite ansätze and for the Cayley transform and we obtain almost guaranteed convergence to the global optimum if the gradient descent terminates.