

# Necessary and sufficient conditions for correctness of complex Langevin

Michael Mandl<sup>1,\*</sup> , Erhard Seiler<sup>2</sup>  and Dénes Sexty<sup>1</sup> 

<sup>1</sup> Institute of Physics, NAWI Graz, University of Graz, Universitätsplatz 5, 8010 Graz, Austria

<sup>2</sup> Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Boltzmannstraße 8, 85748 Garching bei München, Germany

E-mail: [michael.mandl@uni-graz.at](mailto:michael.mandl@uni-graz.at), [ehs@mpp.mpg.de](mailto:ehs@mpp.mpg.de) and [denes.sexy@uni-graz.at](mailto:denes.sexy@uni-graz.at)

Received 8 September 2025; revised 13 November 2025

Accepted for publication 20 November 2025

Published 8 December 2025



CrossMark

## Abstract

We derive a family of correctness conditions for complex Langevin simulations. In particular, we show that if in a given theory the expectation values of all observables within a particular space satisfy the theory's Schwinger–Dyson equations as well as certain bounds, then these expectation values are necessarily correct. In fact, these findings are not only valid in the context of complex Langevin simulations, but they also hold for general probability densities on complex manifolds, given an initial complex density on a real manifold. We stress that, while the proposed conditions are necessary and sufficient in a mathematical sense, their practical use is not to prove the correctness of obtained simulation results. Rather, they are mainly useful for detecting incorrect convergence. In particular, we test these criteria in a few simple one- and two-dimensional toy models and find that they are indeed capable of ruling out incorrect results without the need of exact solutions.

Keywords: sign problem, complex Langevin, Schwinger–Dyson equations, correctness condition

\* Author to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

## 1. Introduction

The numerical sign problem remains among the most challenging problems in many-particle physics to this day. It prevents the numerical evaluation of integrals over highly oscillating functions with a large number of degrees of freedom with standard techniques. Such integrals arise, for instance, in the nonperturbative study of strongly interacting quantum field theories with a nonzero density of fermions by means of first principle lattice simulations. Moreover, they are also prevalent in lattice approaches to the study of real-time quantum field theories, as well as in various other theories of interest in physics.

The complex Langevin method [1, 2] can be used—in principle—in situations where conventional lattice approaches cannot be applied due to the sign problem, thus providing a potential solution. For a theory whose path integral consists of the integration of an oscillatory integrand  $\rho$  over a real manifold  $\mathcal{M}_r$ , the complex Langevin approach makes use of the stochastic evolution in a complexified field space  $\mathcal{M}_c$ . Effectively, it replaces  $\rho$  by a real and nonnegative distribution  $P$  on  $\mathcal{M}_c$ , which thus allows for a probabilistic interpretation similar to the one used in usual Monte-Carlo approaches. The complex Langevin evolution is considered successful if and only if  $P$  is equivalent to  $\rho$  in the sense that both produce the same expectation values on an appropriate space of holomorphic observables. The existence of such a probability density  $P$  for a given  $\rho$  has been established quite generally [3–5]. What is less clear, however, is whether complex Langevin is capable of producing such  $P$  for generic  $\rho$ . In fact, complex Langevin simulations may suffer from the infamous wrong convergence problem, meaning that they might fail to reproduce the correct expectation values, at least without additional modifications.

Establishing the correctness of complex Langevin simulation results is thus of utmost importance in order to ensure the reliability of the method. While there is a plethora of potential correctness criteria to be found in the literature, as we shall discuss, none of them appear to be entirely satisfactory and free of loopholes. In this work, we attempt to fill this gap by deriving a set of conditions that are both necessary and sufficient to ensure that a probability density  $P$  on  $\mathcal{M}_c$  produces the same expectation values as a given complex density  $\rho$  on  $\mathcal{M}_r$ . The correctness criterion proposed here can thus be applied to complex Langevin results, but it is, in fact, more general than that. Indeed, it can be used in any situation where a complex density on a real manifold  $\mathcal{M}_r$  is traded for a probability density on the complexification of  $\mathcal{M}_r$ , see, e.g. [4–7].

This work is structured as follows: after introducing a number of useful definitions and terminology in section 2, we derive the new correctness condition in section 3. In section 4, we introduce the complex Langevin approach, which we then use to test the correctness condition. The results of this are presented in section 5. Finally, in section 6, we provide a few ideas on how to employ the criterion in more realistic models than the ones considered here, before concluding in section 7. The appendices contain some auxiliary results and derivations.

## 2. Generalities

### 2.1. Manifolds and densities

We consider models defined on a real manifold  $\mathcal{M}_r$ , for instance the configuration space of a lattice gauge theory. The manifold  $\mathcal{M}_r$  generally is of the form

$$\mathcal{M}_r = \mathbb{R}^n \times G^m, \tag{1}$$

where  $G$  is a compact Lie group such as  $SU(N)$ . The case  $n = 0$  is interpreted as the absence of the factor  $\mathbb{R}^n$ , while  $m = 0$  is interpreted as the absence of the factor  $G^m$ . On  $\mathcal{M}_r$ , we consider the measure

$$d\mu \equiv d^n x \times d^m g, \tag{2}$$

i.e. the product of the Lebesgue measure  $d^n x$  on  $\mathbb{R}^n$  with the Haar measure  $d^m g$  on  $G^m$ .  $\mathcal{M}_r$  has a complexification to a complex manifold  $\mathcal{M}_c$ ,

$$\mathcal{M}_c = \mathbb{C}^n \times G_c^m, \tag{3}$$

where  $G_c$  is the complexification of  $G$ . We denote the arguments of functions on  $\mathcal{M}_r$  by  $(\mathbf{x}, \mathbf{g})$  and when complexified by  $(\mathbf{z}, \mathbf{g}_c)$ . Occasionally, for notational brevity we will omit the arguments of functions in this work when they are evident from the context.

The models we consider are defined by a complex density  $\rho(\mathbf{x}, \mathbf{g})$ , which has a holomorphic continuation to  $\mathcal{M}_c$ . In many cases,  $\rho$  is given in terms of an action  $S(\mathbf{x}, \mathbf{g})$  via

$$\rho(\mathbf{x}, \mathbf{g}) \propto \exp(-S(\mathbf{x}, \mathbf{g})). \tag{4}$$

We assume  $\rho$  to be normalized, such that

$$\int_{\mathcal{M}_r} \rho d\mu = 1. \tag{5}$$

If  $\mathcal{M}_r$  is noncompact, i.e.  $n > 0$ , we furthermore assume that

$$\rho_m(\mathbf{x}) \equiv \sup_{\mathbf{g} \in G^m} |\rho(\mathbf{x}, \mathbf{g})| \tag{6}$$

goes to zero at least as fast as a Gaussian for  $|\mathbf{x}| \rightarrow \infty$ , i.e.

$$\rho_m(\mathbf{x}) \leq A \exp(-a|\mathbf{x}|^2), \tag{7}$$

with some positive constants  $A$  and  $a$ . This condition could likely be weakened, but it helps to keep the discussion as simple as possible. In any case, the condition is always satisfied in the usual lattice models.

Another requirement that is met for most conventional models, which we also impose here, is that the first partial derivatives of  $\log \rho$  with respect to both the coordinates and the group degrees of freedom (see below) grow at most like a polynomial in  $\mathbf{x}$  for all  $\mathbf{g}$ . Finally, we assume that  $\rho$  does not vanish on  $\mathcal{M}_r$ :

$$\rho(\mathbf{x}, \mathbf{g}) \neq 0 \quad \forall (\mathbf{x}, \mathbf{g}) \in \mathcal{M}_r. \tag{8}$$

This is not really a restriction because by Cauchy's theorem we can always deform the integration manifold  $\mathcal{M}_r$  away from any zeros, for instance by shifting by some vector with small purely imaginary components, without changing the expectation values of observables.

## 2.2. Observables and Schwinger–Dyson operators

Let us consider the (commutative) algebra  $\mathfrak{A}$  of observables generated by the coordinates of  $\mathbb{R}^n$  and the matrix elements<sup>3</sup> of the finite-dimensional representations of  $G^m$  (note that in the case of  $G = \text{SU}(N)$  we only require the matrix elements of the fundamental representation as generators of the algebra). All the elements  $f \in \mathfrak{A}$  have analytic continuations to  $\mathcal{M}_c$ , which we denote by the same letter  $f$ . The constant function with value 1 is the unit element of  $\mathfrak{A}$ , denoted by  $\mathbb{1}$ . It turns out, however, that we actually require a somewhat larger space of observables for our purposes, as will be explained below.

It is well known that the correct expectation values of observables have to satisfy the so-called Schwinger–Dyson equations (SDEs), which are infinitely many identities characterizing  $\rho$ . The Schwinger–Dyson (SD) operators are differential operators acting on the observables, but—in general—they do not map  $\mathfrak{A}$  into itself. To describe the SDEs in this general setting, we first need to establish a few definitions.

Since the functions  $f \in \mathfrak{A}$  depend on two kinds of variables,  $\mathbf{x}$  and  $\mathbf{g}$ , we also require two kinds of derivatives. The first kind are simply the partial derivatives  $\partial_i$  with respect to the coordinates  $x_i$  of  $\mathbb{R}^n$  or, after complexification, the coordinates  $z_i$  of  $\mathbb{C}^n$ . The second ones are derivatives with respect to the group variables. They are determined by choosing a basis

$$\{X_j | j = 1, \dots, d\}, \quad d = \dim(G) \tag{9}$$

of the Lie algebra  $\mathfrak{g}$  of  $G$  and, correspondingly, a basis of the Lie algebra  $\mathfrak{g}^m$  as

$$\{X_{j,k} | j = 1, \dots, d; k = 1, \dots, m\}. \tag{10}$$

With this, we define the derivations (usually referred to as left-derivatives)

$$D_{j,k}f(\mathbf{x}, \mathbf{g}) \equiv \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (f(\mathbf{x}, \exp(\epsilon X_{j,k}) \mathbf{g}) - f(\mathbf{x}, \mathbf{g})) \tag{11}$$

and extend their action to the analytic continuation of the functions  $f \in \mathfrak{A}$ . Corresponding to the two kinds of partial derivatives, there are two kinds of SD operators:

$$A_i = \partial_i + \frac{\partial_i \rho}{\rho} \tag{12}$$

and

$$A_{j,k} = D_{j,k} + \frac{D_{j,k} \rho}{\rho}. \tag{13}$$

**Remark.** The SD operators map  $\mathfrak{A}$  into itself only in the case that the partial derivatives of  $\log \rho$  are polynomials. More precisely, if  $\rho$  vanishes somewhere on  $\mathcal{M}_c$ , the action of the SD operators on  $\mathfrak{A}$  will generically produce poles at the zeros of  $\rho$ . For this reason, as in [8], we are forced to enlarge  $\mathfrak{A}$  to

$$\mathcal{H} \equiv \mathfrak{A} + \sum_i A_i \mathfrak{A} + \sum_{j,k} A_{j,k} \mathfrak{A}. \tag{14}$$

<sup>3</sup> Notice that the matrix elements of  $\mathbf{g} \in G$  are functions of  $\mathbf{g}$ , such that the observables in  $\mathfrak{A}$  are really functions on  $\mathcal{M}_r$ .

It is not hard to see that all  $f \in \mathcal{H}$  are integrable with respect to  $|\rho|d\mu$ . In section 5.2, we shall discuss a specific example in which  $\mathfrak{A} \neq \mathcal{H}$ .

Now, the ‘correct expectation’ of a function  $f \in \mathcal{H}$  is defined by

$$\langle f \rangle_\rho \equiv \int_{\mathcal{M}_r} f \rho d\mu, \tag{15}$$

where

$$\langle \mathbb{1} \rangle_\rho = 1 \tag{16}$$

due to the normalization of  $\rho$ . Indeed,  $\langle f \rangle_\rho$  is invariant under continuous deformations or shifts of  $\mathcal{M}_r$ , owing to the holomorphicity of both  $\rho$  and  $f$ .

### 2.3. Correctness of probability measures on $\mathcal{M}_c$

Let  $dP$  be a probability measure on  $\mathcal{M}_c$ , which may (but does not have to) be the equilibrium measure of a complex Langevin process with or without a kernel (see section 4).

**Definition.** We say that the probability measure  $dP$  is *correct* if and only if

$$\int_{\mathcal{M}_c} f dP = \langle f \rangle_\rho \quad \forall f \in \mathcal{H}. \tag{17}$$

**Remark.** If  $\rho$  vanishes somewhere on  $\mathcal{M}_c$ ,  $dP$  can only be correct if it also vanishes there, because otherwise the singularities in  $f \in \mathcal{H}$  would cause the left-hand side of (17) to diverge.

We factorize  $\rho$  as

$$\rho(\mathbf{x}, \mathbf{g}) = w(\mathbf{x}) \rho_r(\mathbf{x}, \mathbf{g}), \tag{18}$$

such that both factors are continuous,  $w > 0$ , and

$$\|f\|_w \equiv \sup_{\mathcal{M}_r} |f(\mathbf{x}, \mathbf{g})| w(\mathbf{x}) < \infty \quad \forall f \in \mathcal{H}, \tag{19}$$

as well as

$$C \equiv \int_{\mathcal{M}_r} |\rho_r(\mathbf{x}, \mathbf{g})| d\mu < \infty. \tag{20}$$

Note that the factor  $w$  is independent of the group variables  $\mathbf{g}$ .

To be clear, since we assume the partial derivatives of  $\log \rho$  to grow at most polynomially, these conditions require that  $w(\mathbf{x})$  decays faster than any power of  $|\mathbf{x}|$ , while (20) must hold at the same time. In fact, for reasons stated below, we actually require  $w(\mathbf{x})$  to decay at least like a Gaussian, i.e.

$$w(\mathbf{x}) \leq B \exp(-b|\mathbf{x}|^2), \tag{21}$$

with positive constants  $B$  and  $b$ , similar to (7). A simple choice for  $w$  is

$$w(\mathbf{x}) \equiv \rho_m(\mathbf{x})^\alpha, \quad \rho_r(\mathbf{x}, \mathbf{g}) \equiv \rho(\mathbf{x}, \mathbf{g}) / w(\mathbf{x}), \quad 0 < \alpha < 1, \tag{22}$$

where  $\rho_m$  was defined in (6). Note that this implies

$$\sup_{\mathbf{g} \in G^m} |\rho_r(\mathbf{x}, \mathbf{g})| = w(\mathbf{x})^{1/\alpha-1}. \tag{23}$$

With this choice,  $w(\mathbf{x})$  indeed decays like a Gaussian due to (7) and is continuous, as is shown in appendix A.

### 3. Correctness conditions

With the definitions introduced in the previous section, we are now in the position to formulate two conditions, the combination of which is both necessary and sufficient for ensuring the correctness of the probability measure  $dP$ . We proceed by proving the necessity and sufficiency in turn.

#### 3.1. Two necessary conditions for correctness

**Condition 1.** If  $dP$  is correct, i.e. (17) holds, it satisfies the SDEs

$$\int_{\mathcal{M}_c} (A_i f) dP = 0, \quad i = 1, \dots, n \quad \forall f \in \mathfrak{A} \tag{24}$$

and

$$\int_{\mathcal{M}_c} (A_{j,k} f) dP = 0, \quad j = 1, \dots, d; \quad k = 1, \dots, m \quad \forall f \in \mathfrak{A}. \tag{25}$$

To prove (24) and (25), we only have to integrate the right-hand side of (17) (i.e. (15)) by parts, which under our assumptions can be carried out without the appearance of boundary terms, and use the definitions of  $A_i$  and  $A_{j,k}$  in (12) and (13). Note that the validity of (17) for all  $f \in \mathcal{H}$  implies the validity of (24) and (25) only for  $f \in \mathfrak{A}$ .

The second condition is another straightforward consequence of (17), using the norm  $\|f\|_w$  in (19) and the constant  $C$  defined in (20):

**Condition 2.** If  $dP$  is correct, i.e. it satisfies (17), the following bounds hold for all  $f \in \mathcal{H}$ :

$$\left| \int_{\mathcal{M}_c} f dP \right| \leq C \|f\|_w. \tag{26}$$

**Proof.** Using (17) and (18), we have

$$\left| \int_{\mathcal{M}_c} f(\mathbf{z}, \mathbf{g}_c) dP \right| = \left| \int_{\mathcal{M}_r} f(\mathbf{x}, \mathbf{g}) w(\mathbf{x}) \rho_r(\mathbf{x}, \mathbf{g}) d\mu \right|, \tag{27}$$

for which we can find a trivial upper bound, namely

$$\left| \int_{\mathcal{M}_c} f(\mathbf{z}, \mathbf{g}_c) dP \right| \leq \left( \int_{\mathcal{M}_r} |\rho_r(\mathbf{x}, \mathbf{g})| d\mu \right) \sup_{\mathcal{M}_r} |f(\mathbf{x}, \mathbf{g})| w(\mathbf{x}) = C \|f\|_w, \tag{28}$$

which is precisely (26). □

### 3.2. Sufficiency of the necessary conditions

We now go on to show that conditions 1 and 2 in combination are also sufficient to guarantee correctness, based on our earlier assumption that  $\rho$  does not vanish anywhere on  $\mathcal{M}_r$ . To this end, let us assume that (26) holds. This means that the linear functional  $T_P$  on  $\mathcal{H}$ , defined by

$$(T_P, f) \equiv \int_{\mathcal{M}_c} f dP, \quad (29)$$

is bounded in the norm  $\|\cdot\|_w$ , or, equivalently, that the linear functional  $\tilde{T}_P$  on  $\tilde{\mathcal{H}} \equiv w\mathcal{H}$ , defined by

$$(\tilde{T}_P, \tilde{f}) \equiv (T_P, f) \quad \text{with} \quad \tilde{f} = wf, \quad (30)$$

satisfies

$$|(\tilde{T}_P, \tilde{f})| \leq C \|\tilde{f}\|_\infty \quad \forall \tilde{f} \in \tilde{\mathcal{H}}, \quad \|\tilde{f}\|_\infty \equiv \sup_{\mathcal{M}_r} |\tilde{f}(\mathbf{x}, \mathbf{g})|. \quad (31)$$

Note that  $\tilde{\mathcal{H}} \subset \mathfrak{C}_0(\mathcal{M}_r)$ , where  $\mathfrak{C}_0(\mathcal{M}_r)$  is the space of all continuous functions on  $\mathcal{M}_r$  vanishing at infinity. This is due to our requirements that  $\rho$  does not vanish on  $\mathcal{M}_r$  and that any possible large growth of  $f$  at infinity is counteracted by the fast decay of  $w$ . For the latter to be true, it was necessary to restrict the growth of the partial derivatives of  $\log \rho$  earlier.

A crucial fact is the following

**Proposition 1.** *Any linear functional  $\tilde{T}$  on  $\tilde{\mathcal{H}} \subset \mathfrak{C}_0(\mathcal{M}_r)$  that is bounded as in (31) can be extended to a linear functional on all of  $\mathfrak{C}_0(\mathcal{M}_r)$  with the same bound.*

**Proof.** This follows directly from the *Hahn–Banach* theorem; see, e.g. [9] (theorem 5.7) or [10] (theorem 3.3).  $\square$

The extension given by this theorem, which we denote by the same symbol  $\tilde{T}$ , may not be unique, but we will in the end only be interested in its action on  $\tilde{\mathcal{H}}$  anyway. The usefulness of this extension comes from another abstract result, known as the *Riesz–Markov* theorem; see, e.g. [11] (theorem 6.19). Adapting the notation to our case, it says

**Theorem (Riesz–Markov).** *If  $X$  is a locally compact Hausdorff space, then every bounded linear functional  $\tilde{T}$  on  $\mathfrak{C}_0(X)$  is represented by a unique regular complex Borel measure  $\nu$  in the sense that  $(\tilde{T}, \tilde{f}) = \int_X \tilde{f} d\nu$  for every  $\tilde{f} \in \mathfrak{C}_0(X)$ . Moreover,  $\nu$  is of finite total variation, i.e.  $\|\tilde{T}\| \equiv \int_X d|\nu| < \infty$ .*

In our context,  $\mathcal{M}_r$  is locally compact as it is of finite dimension by assumption. Also note that a Borel measure  $d\nu$  on  $\mathcal{M}_r$  always represents a generalized function (actually a tempered distribution, see [12])  $\tilde{\sigma}$ , so we can write

$$d\nu = \tilde{\sigma}(\mathbf{x}, \mathbf{g}) d\mu. \quad (32)$$

Hence, we obtain

**Proposition 2.** *Any bounded linear functional  $\tilde{T}$  on  $\mathfrak{C}_0(\mathcal{M}_r)$  is given by a complex measure, which we represent by a distributional density  $\tilde{\sigma}$  on  $\mathcal{M}_r$  as follows:*

$$(\tilde{T}, \tilde{f}) = \int_{\mathcal{M}_r} \tilde{\sigma}(\mathbf{x}, \mathbf{g}) \tilde{f}(\mathbf{x}, \mathbf{g}) d\mu, \quad \tilde{f} \in \mathfrak{C}_0(\mathcal{M}_r), \quad (33)$$

where  $\tilde{\sigma}$  satisfies

$$\int_{\mathcal{M}_r} |\tilde{\sigma}| d\mu < \infty. \quad (34)$$

Restricting to the subset  $\tilde{\mathcal{H}} \subset \mathcal{C}_0(\mathcal{M}_r)$  and applying this proposition to the functional  $\tilde{T}_P$  that is generated by the probability measure  $dP$  according to (30) and (29) and that we have just shown to be bounded, we find that it is represented by a complex measure  $\tilde{\sigma}_P$  of bounded total variation, i.e.

$$(\tilde{T}_P, \tilde{f}) = \int_{\mathcal{M}_r} \tilde{\sigma}_P(\mathbf{x}, \mathbf{g}) \tilde{f}(\mathbf{x}, \mathbf{g}) d\mu, \quad \tilde{f} = wf, \quad f \in \mathcal{H}. \quad (35)$$

To sum up, we have obtained

**Proposition 3.** *Let  $dP$  be a probability measure on  $\mathcal{M}_c$ , satisfying for all  $f \in \mathcal{H}$*

$$\left| \int_{\mathcal{M}_c} f(\mathbf{z}, \mathbf{g}_c) dP \right| \leq C \|f\|_w, \quad (36)$$

with  $\|f\|_w$  and  $C$  defined in (19) and (20), respectively, then, again for all  $f \in \mathcal{H}$ , we have

$$\int_{\mathcal{M}_c} f(\mathbf{z}, \mathbf{g}_c) dP = \int_{\mathcal{M}_r} \tilde{\sigma}_P(\mathbf{x}, \mathbf{g}) w(\mathbf{x}) f(\mathbf{x}, \mathbf{g}) d\mu, \quad (37)$$

with  $\tilde{\sigma}_P$  of finite total variation.

Defining

$$\sigma_P(\mathbf{x}, \mathbf{g}) \equiv w(\mathbf{x}) \tilde{\sigma}_P(\mathbf{x}, \mathbf{g}), \quad (38)$$

this can be rewritten as

$$\int_{\mathcal{M}_c} f(\mathbf{z}, \mathbf{g}_c) dP = \int_{\mathcal{M}_r} \sigma_P(\mathbf{x}, \mathbf{g}) f(\mathbf{x}, \mathbf{g}) d\mu. \quad (39)$$

In other words, if the probability measure  $dP$  on  $\mathcal{M}_c$  is tested only against functions in  $\mathcal{H}$ , it can be replaced by  $\sigma_P$ , which is concentrated on  $\mathcal{M}_r$ . The density  $\sigma_P$  decays at least like a Gaussian in  $|\mathbf{x}|$  (since  $w(\mathbf{x})$  has the same property and  $\tilde{\sigma}_P$  is of finite total variation) for all  $\mathbf{g}$ . We now show that condition 1 (i.e. the SDEs (24) and (25)) together with condition 2 (i.e. the bounds (26)) implies correctness.

**Proof.** Integration by parts of (24) and (25), after using (39), rolls the respective SD operators over to  $\sigma_P$ , changing them into their transposes as

$$\int_{\mathcal{M}_r} f(\mathbf{x}, \mathbf{g}) A_i^T \sigma_P(\mathbf{x}, \mathbf{g}) d\mu = 0 \quad \forall f \in \mathfrak{A}, \quad i = 1, \dots, n \quad (40)$$

and

$$\int_{\mathcal{M}_r} f(\mathbf{x}, \mathbf{g}) A_{j,k}^T \sigma_P(\mathbf{x}, \mathbf{g}) d\mu = 0 \quad \forall f \in \mathfrak{A}, \quad j = 1, \dots, d; k = 1, \dots, m, \quad (41)$$

where

$$A_i^T = -\partial_i + \frac{\partial_i \rho}{\rho} \quad \text{and} \quad A_{j,k}^T = -D_{j,k} + \frac{D_{j,k} \rho}{\rho}. \quad (42)$$

Note that under our conditions, there are once again no boundary terms here. Up to now, we have treated  $\sigma_P$  as a distribution, in fact one of order zero [10] according to the bounds (26).  $\sigma_P$  therefore represents a complex (Borel) measure by the Riesz–Markov theorem. But we have not yet addressed the question whether it is a differentiable function, so the derivatives in (40) and (41) have to be understood in the distributional sense.

We would like to conclude from (40) and (41) that (for all  $i, j$ , and  $k$ )

$$A_i^T \sigma_P = 0 \quad \text{and} \quad A_{j,k}^T \sigma_P = 0. \quad (43)$$

These differential equations then would imply that  $\sigma_P$  is really a differentiable function. Indeed, this will turn out to be true because there are enough functions in  $\mathfrak{A}$  such that a distribution  $T$  satisfying  $(T, f) = 0 \forall f \in \mathfrak{A}$  vanishes identically. However, in order to see this more explicitly, we require a few more arguments.

We consider functions of the form

$$f_{\mathbf{k}}(\mathbf{x}, \mathbf{g}) = \exp(\mathbf{i}\mathbf{k} \cdot \mathbf{x}) h(\mathbf{g}), \quad (44)$$

where  $h$  is in the sub-algebra  $\mathfrak{A}_G$  of  $\mathfrak{A}$  consisting of functions that do not depend on  $\mathbf{x}$ . Note that the functions  $f_{\mathbf{k}}$  are not in  $\mathfrak{A}$ , but they are contained in its closure  $\overline{\mathfrak{A}}$  with respect to  $\|\cdot\|_w$ , as is shown in appendix B.

Now, to prove (43), we consider the functions

$$\hat{F}_i(\mathbf{k}) = \int_{\mathcal{M}_r} \sigma_P(\mathbf{x}, \mathbf{g}) A_i f_{\mathbf{k}}(\mathbf{x}, \mathbf{g}) \, d\mu \quad (45)$$

and

$$\hat{F}_{j,k}(\mathbf{k}) = \int_{\mathcal{M}_r} \sigma_P(\mathbf{x}, \mathbf{g}) A_{j,k} f_{\mathbf{k}}(\mathbf{x}, \mathbf{g}) \, d\mu, \quad (46)$$

which all vanish due to (40) and (41), and the fact that  $f_{\mathbf{k}}$  is in  $\overline{\mathfrak{A}}$ . Moreover,  $\hat{F}_i(\mathbf{k})$  is the Fourier transform of

$$F_i(\mathbf{x}) = \int_{G^m} h(\mathbf{g}) A_i^T \sigma_P(\mathbf{x}, \mathbf{g}) \, d^m g, \quad (47)$$

and  $\hat{F}_{j,k}(\mathbf{k})$  the Fourier transform of

$$F_{j,k}(\mathbf{x}) = \int_{G^m} h(\mathbf{g}) A_{j,k}^T \sigma_P(\mathbf{x}, \mathbf{g}) \, d^m g, \quad (48)$$

which also all vanish because the Fourier transformation is one-to-one on the space of (tempered) distributions on  $\mathbb{R}^n$  [12]. By the Peter–Weyl theorem (see, e.g. [13] (theorem 4.1) or [14] (theorem 1.12)),  $\mathfrak{A}_G$  is dense in the space  $\mathfrak{C}(G^m)$  of continuous functions on  $G^m$ . Therefore, it is also dense in the space  $\mathfrak{C}^\infty(G^m)$ , the standard space of test functions on  $G^m$ , in the appropriate topology. As a consequence, the vanishing of (47) and (48) for arbitrary  $h(\mathbf{g})$  implies (43).

Owing to the requirement that  $\rho$  does not vanish anywhere on  $\mathcal{M}_r$ , the differential equations (43) have the general solution

$$\sigma_P(\mathbf{x}, \mathbf{g}) = c\rho(\mathbf{x}, \mathbf{g}) \tag{49}$$

with  $c \in \mathbb{C}$ . By choosing  $f = \mathbb{1}$  and from the fact that  $P$  is normalized, we find that  $c = 1$  and correctness (17) holds.  $\square$

Let us summarize what we have shown:

**Theorem 1.** *Let  $P$  be a probability distribution on  $\mathcal{M}_c$ . Assume that  $\rho$  does not vanish on  $\mathcal{M}_r$  and that all partial derivatives of  $\log \rho$  grow at most like a polynomial for  $|\mathbf{x}| \rightarrow \infty$ . Then, the following are equivalent:*

(a)  $P$  is correct in the sense that for any  $f \in \mathcal{H}$

$$\int_{\mathcal{M}_c} f dP = \int_{\mathcal{M}_r} f \rho d\mu. \tag{50}$$

(b)  $P$  satisfies the SDEs (24) and (25) as well as the bounds

$$\left| \int_{\mathcal{M}_c} f dP \right| \leq C \|f\|_w, \tag{51}$$

with  $\rho(\mathbf{x}, \mathbf{g}) = w(\mathbf{x})\rho_r(\mathbf{x}, \mathbf{g})$ ,  $w$  obeying (21),

$$\|f\|_w \equiv \sup_{\mathcal{M}_r} |f(\mathbf{x}, \mathbf{g})| w(\mathbf{x}) < \infty \quad \forall f \in \mathcal{H}, \tag{52}$$

and

$$C \equiv \int_{\mathcal{M}_r} |\rho_r(\mathbf{x}, \mathbf{g})| d\mu < \infty. \tag{53}$$

**Remark.** Up to now, we have tried to keep things as general as possible. As a consequence, theorem 1 does not provide one single condition for correctness, but in fact an entire family of conditions: For one, the choice of the real manifold  $\mathcal{M}_r$  is not unique, as, for instance, continuous deformations (with fixed endpoints) of  $\mathcal{M}_r$  give rise to the same ‘correct expectations’ (15) due to Cauchy’s theorem and the assumed holomorphicity of  $f$  and  $\rho$ . However, the bounds  $C\|f\|_w$  do depend on the choice of  $\mathcal{M}_r$  (keep in mind, though, that  $\mathcal{M}_r$  must always be chosen such that it contains no zeros of  $\rho$ ). In addition, the decomposition of  $\rho$  into  $w$  and  $\rho_r$  in (18) is also arbitrary as long as the resulting  $C$  and  $\|f\|_w$  ( $\forall f \in \mathcal{H}$ ) are finite and  $w$  decays at least like a Gaussian. The constant  $C$  and the norm  $\|f\|_w$ , as well as their product, will, of course, differ for different decompositions.

What is perhaps even more important is the freedom in the choice of norms. We have used in (19) and (20) the supremum norm for the observables and the  $L^1$  norm for the reduced density  $\rho_r$ . One may, however, just as well choose, for instance,  $L^p$  and  $L^q$  norms with  $1/p + 1/q = 1$  instead. In particular, one could choose  $p = q = 2$ , i.e.  $L^2$  norms for both quantities, replacing (19) and (20) by

$$\|f\|_w \equiv \sqrt{\int_{\mathcal{M}_r} |f(\mathbf{x}, \mathbf{g})|^2 w(\mathbf{x})^2 d\mu} < \infty \quad \forall f \in \mathcal{H} \tag{54}$$

and

$$C \equiv \sqrt{\int_{\mathcal{M}_r} |\rho_r(\mathbf{x}, \mathbf{g})|^2 d\mu} < \infty, \quad (55)$$

respectively. It should be understood that the practical applicability of the proposed correctness criterion might strongly depend on all of these choices. A detailed analysis systematically comparing different possible correctness conditions arising from our proposal is, however, beyond the scope of this work.

#### 4. Basics of complex Langevin simulations

Having developed a novel correctness criterion for probability distributions  $P$  on complexified manifolds  $\mathcal{M}_c$ , a logical next step is to put it to the test in practical applications. Our main interest is its applicability in the context of complex Langevin simulations, which we shall thus briefly motivate and introduce in the present section.

##### 4.1. The sign problem

In the path integral picture of quantum field theory, expectation values of observables  $\mathcal{O}$  are defined as integrals of the form

$$\langle \mathcal{O} \rangle_\rho \equiv \int \mathcal{D}\Phi \mathcal{O}[\Phi] \rho[\Phi], \quad (56)$$

where  $\Phi$  collectively denotes the dynamical degrees of freedom of the theory under consideration,  $\mathcal{D}\Phi$  is the appropriate path integral measure for these fields and  $\rho[\Phi]$  is some density giving a nontrivial weight to each field configuration. Typically,

$$\rho[\Phi] = \frac{e^{iS_M[\Phi]}}{\int \mathcal{D}\Phi e^{iS_M[\Phi]}} \quad \text{or} \quad \rho[\Phi] = \frac{e^{-S_E[\Phi]}}{\int \mathcal{D}\Phi e^{-S_E[\Phi]}}, \quad (57)$$

for a theory with an action  $S_M[\Phi]$  in Minkowski spacetime, or, (after performing a suitable Wick rotation) an action  $S_E[\Phi]$  in Euclidean space, respectively. The normalization factors in the denominator ensure that  $\langle \mathbb{1} \rangle_\rho = 1$ .

The most successful non-perturbative tool for computing  $\langle \mathcal{O} \rangle$  from first principles is lattice quantum field theory, in which—conventionally—the integral in (56) is approximated by a finite sum over configurations  $\{\Phi\}_i$  distributed according to  $\rho[\Phi]$ , which is the traditional importance sampling approach. For this to make sense mathematically, however, one requires that  $\rho[\Phi]$  is real and nonnegative, i.e. that it can be interpreted as a probability density. Clearly, this is not the case in Minkowski spacetime, which is why in the vast majority of lattice approaches one generically works in Euclidean space. However, the Euclidean action  $S_E$  is not guaranteed to be real either. Cases in which  $S_E$  acquires a nonzero imaginary part include (but are not restricted to) quantum field theories (in thermodynamic equilibrium) at nonvanishing chemical potential or with topological terms.

In all of the aforementioned cases, the straightforward application of usual lattice techniques based on importance sampling fails. This is the infamous sign (or complex-action) problem [15]; see, e.g. [16–18] for reviews.

#### 4.2. The complex Langevin equation

One particular method that is aimed at circumventing the sign problem is the complex Langevin approach [1, 2] based on stochastic quantization [19]. The underlying mechanism is the stochastic evolution of complexified field degrees of freedom in a fictitious time dimension  $\tau$ , governed by the complex Langevin equation. For simplicity, let us restrict to  $N$  real scalar degrees of freedom  $\Phi_i^R$  ( $i = 1, \dots, N$ ) described by the Euclidean<sup>4</sup> action  $S[\Phi^R]$ . Then, the complex Langevin equation reads<sup>5</sup>

$$d\Phi_i(\tau) = D_i(\Phi) d\tau + H_{ij} dw_j(\tau), \quad (58)$$

where the  $\Phi_i \equiv \Phi_i^R + i\Phi_i^I$  result from the continuation of each  $\Phi_i^R$  into the complex plane. Moreover,  $D_i$  denotes the so-called drift term,

$$D_i = -H_{ik} H_{jk} \frac{\partial S}{\partial \Phi_j} + \frac{\partial (H_{ik} H_{jk})}{\partial \Phi_j}, \quad (59)$$

$H_{ij}$  is a (generally  $\Phi$ -dependent) complex  $N \times N$  matrix called the kernel and  $dw_i$  describes a standard Wiener process in each degree of freedom, such that

$$\langle dw_j(\tau) \rangle = 0, \quad \langle dw_j(\tau) dw_k(\tau) \rangle = 2\delta_{jk} d\tau \quad (60)$$

for all field indices  $j$  and  $k$ . In order for this approach to be well defined, the action must allow for an analytic continuation  $S[\Phi]$  in its arguments and the same must be true for the observables  $\mathcal{O}[\Phi]$ . We also mention that the extension of the Langevin equation to gauge theories is relatively straightforward; see, e.g. [20].

Now, the stochastic evolution governed by the complex Langevin equation gives rise to a (real and nonnegative) probability distribution  $P(\Phi_i^R, \Phi_i^I; \tau)$  on the complex manifold, whose  $\tau$ -dependence is, in turn, determined by an associated so-called Fokker–Planck equation. The exact form of the latter, however, is irrelevant for the present discussion. In the following, we assume that  $P$  has an equilibrium limit

$$P(\Phi_i^R, \Phi_i^I) \equiv \lim_{\tau \rightarrow \infty} P(\Phi_i^R, \Phi_i^I, \tau). \quad (61)$$

Note that this assumption is not always justified, but only in the case when there exists such a limit is the complex Langevin approach applicable. As a consequence, the complex Langevin evolution after equilibration indeed leaves us with a probability distribution  $P$  on a complex manifold  $\mathcal{M}_c$ , for which the correctness condition developed in section 3 can be checked. Before doing so below in the context of a few simple models, however, let us first recall possible reasons for why and how complex Langevin simulations may fail in the first place.

#### 4.3. Reasons of failure for the complex Langevin approach

The study of possible error sources for complex Langevin simulations has a rather long history [8, 21–39]. One common issue is the presence of so-called runaway trajectories, i.e. paths in

<sup>4</sup> Here and in the following, we work in Euclidean space and suppress the subscript ‘E’ on the Euclidean action. Note, however, that the discussion can straightforwardly be transferred to Minkowski spacetime via the replacement  $-S_E \rightarrow iS_M$ .

<sup>5</sup> Implicit summation over indices occurring twice in a product is assumed throughout.

complexified field space along which the classical evolution (i.e. the one excluding the noise term in (58)) diverges. Their existence may potentially lead to large discretization effects or even cause simulations to crash when the involved drift terms become too large. However, the effects of runaway trajectories can often be mitigated or even eliminated entirely in practice by the use of adaptive step-size control [26], implicit solvers [40], and (in the case of gauge theories) the gauge cooling technique [41]. At the very least, it is usually possible to detect whether or not runaways have an impact on a given simulation by, for instance, monitoring the distance of the simulation trajectory from the real manifold.

In addition, complex Langevin simulations may suffer from a loss of ergodicity, especially in the presence of poles in the drift term. In this case, one may find that the sampled configuration space is split into different disconnected parts [8, 23, 34, 42]. While transitions between these regions are possible in simulations due to the finiteness of the discrete Langevin step size, they become increasingly rare in the continuum limit and their absence can lead to incorrect results. It is not entirely clear how to deal with a loss of ergodicity in the general case, even though in a few models there exist solutions employing, for instance, a kernel or complex noise [34].

Another issue concerns the decay properties of the probability distribution  $P$  resulting from the complex Langevin evolution. Generically, one expects the complex Langevin results to be correct only if  $P$  decays fast enough in all complex directions. More precisely, a correctness criterion for the complex Langevin evolution was formulated in [25, 28] that relies on the sufficiently fast decay of  $P\mathcal{O}$  ( $\mathcal{O}$ , as before, denoting an observable) such that one may perform an integration by parts without the appearance of boundary terms. Most generally, however, one cannot expect such boundary terms to vanish, which is why it is advisable to measure them in practice [35, 36]. If the boundary terms are nonzero, the obtained simulation results are generally incorrect (even though for some observables the expectation values may still come out right).

The converse, however, is not always true, i.e. the absence of boundary terms does not automatically imply the correctness of results [8, 34, 43]. Similarly, while the validity of the SDEs is necessary for correct convergence, it is not sufficient. In particular, as was shown analytically for one-dimensional systems in [8], any linear functional  $T$  (such as the one provided by the complex Langevin evolution) on a suitable space of test functions (such as the algebra  $\mathfrak{A}$  defined in section 2) that obeys the SDEs is given by a linear combination of the form

$$(T;f) = \sum_{i=1}^{n_\gamma} a_i (T_{\gamma_i};f). \quad (62)$$

Here, the  $a_i$  are complex coefficients and

$$(T_{\gamma_i};f) \equiv \int_{\gamma_i} dz f(z) \rho(z) \quad (63)$$

are the ‘expectation values’ of  $f$  along paths  $\gamma_i$ , which are linearly independent so-called integration cycles (in the terminology of [44]). Concretely, an integration cycle  $\gamma_i$  is either a path that connects two distinct zeros of  $\rho(z)$  or a closed noncontractible path. It is important to note that by an integration cycle one really means an equivalence class of integration paths that can be deformed into one another or are equivalent via Cauchy’s theorem. In (62), we have denoted the total number of linearly independent integration cycles by  $n_\gamma$ . Now, if we define (without loss of generality)  $\gamma_1 = \mathcal{M}_r$ , we say that results are correct if and only if  $a_i = \delta_{i1}$ .

The role of integration cycles in the context of complex Langevin simulations was studied in detail in [43], where evidence for the validity of (62) beyond one dimension was established as well. We can summarize the situation as follows: if for a particular complex Langevin simulation boundary terms vanish/the SDEs hold, but the results are nonetheless incorrect, this is necessarily due to contributions from unwanted integration cycles, i.e. from  $\gamma_i \neq \mathcal{M}_r$ . The relationship between vanishing boundary terms, the validity of the SDEs and the so-called complex Langevin consistency conditions [28] is discussed in some detail in appendix C. We mention in passing that the sampling of unwanted integration cycles can in some cases be related to the aforementioned ergodicity problems [8, 34]. Either way, the correctness criterion we propose in this work should—at least in principle—be able to fill the gap left open by the boundary-term (and equivalent) analyses, enabling one to detect incorrect convergence even when boundary terms are zero.

Let us now briefly address the scope of applicability of this correctness criterion in practice. Needless to say, there is little hope in trying to use it to establish unambiguously that a particular simulation gives rise to correct results. Since this would require measuring every observable contained in the space  $\mathfrak{A}$  (or  $\mathcal{H}$ ), which is of infinite size, it is clearly an impossible task. What it should allow one to do, however, is rule out certain results despite the absence of boundary terms or in cases where boundary terms are very small or noisy. To exemplify this idea, below the criterion will be applied in the context of a few simple models for which complex Langevin simulations are known to produce no boundary terms but whose results nonetheless disagree with exact solutions [8, 34, 43]. Obviously, in most realistic systems no exact results are available, making any sort of correctness criterion all the more desirable.

## 5. Examples and numerical checks

In this section, we perform numerical tests on the new condition for correctness for a handful of simple systems with known results. We restrict ourselves to models with  $m = 0$ , i.e. there are no Lie-group-valued degrees of freedom, such that  $\mathcal{M}_r = \mathbb{R}^n$ . A more detailed analysis in the case of gauge theories is deferred to future work. Following the FAIR guiding principles [45], we make the simulation data underlying the results presented in this section [46], as well as the used analysis scripts [47], available online. Moreover, our simulation code is available upon request.

### 5.1. Regular models

We begin by considering models whose densities are given by  $\rho(\mathbf{x}) = e^{-S(\mathbf{x})}$  with polynomial  $S(\mathbf{x})$ . This simplifies the discussion as in this case  $\mathfrak{A} = \mathcal{H}$ , i.e. we can restrict to an algebra of polynomial observables. The numerical simulation and data analysis setup we use for these models is outlined in detail in [43]. In fact, we re-use parts of the data set generated for [43] in the present work.

**5.1.1. One dimension.** The first model to be investigated is given by the action

$$S(x) = \frac{\lambda}{4} x^4, \quad (64)$$

with a single real degree of freedom  $x$  and a complex parameter  $\lambda$  obeying  $|\lambda| = 1$ . The correctness of the probability distribution  $P$  produced by complex Langevin simulations including a kernel in this model was first investigated in [48]. More recently, in [43, 49] the model was

revisited and the connection between the kernel and the relevant integration cycles was investigated. In particular, it was discovered that a kernel can—to a certain extent—be used to control which integration cycles are being sampled in a simulation. Here, we once again revisit this simple model as a testbed for our correctness criterion.

The complex Langevin equation (58) for the theory (64), after complexification  $x \rightarrow z = x + iy$ , simplifies to

$$dz(\tau) = -H^2 S'(z) d\tau + Hd w(\tau), \tag{65}$$

where we have assumed the kernel  $H$  to be independent of  $z$  for simplicity. It is thus merely a constant complex number that we may choose at will. As in [43, 48, 49], we parameterize the kernel<sup>6</sup> with an integer  $m_H$  as

$$H = e^{-i\frac{m_H}{48}}. \tag{66}$$

An important finding of [43, 49] was that by an appropriate choice of kernel one can indeed obtain  $a_i = \delta_{i1}$  in (62), where  $\gamma_1 = \mathcal{M}_r$ . If any cycle other than  $\mathcal{M}_r$  were to give a sizable contribution to the results, the latter would be considered incorrect. This incorrectness, however, could not be detected simply via a boundary-term analysis. In the following, however, we shall demonstrate that it can be detected via the new criterion.

To this end, we make use of the freedom of choosing  $\mathcal{M}_r$  at our discretion. In particular, since the real  $x$  axis does not correspond to a valid integration cycle if  $\text{Re}(\lambda) < 0$ , we define as  $\mathcal{M}_r$  the integration contour resulting from rotating the real axis as  $x = \lambda^{-1/4}\xi$ , along which the analytically continued action  $S(\xi)$  is real. We mention, however, that in a realistic theory the choice of  $\mathcal{M}_r$  is dictated by the physics and one only has the freedom of continuously deforming it, contrary to what is done here. As is detailed in [43], there are two more independent integration cycles in this model. In analogy to [43], we choose the imaginary axis and a contour connecting  $z = i\infty$  with  $z = \infty$ , both after the same rotation as for the real axis. In the following, we set  $\lambda = e^{5i\pi/6}$ . Moreover, from now on we will exclusively use the letter  $z$  to denote the dynamical degree of freedom to avoid possible confusion arising from jumping back and forth between  $x$  and  $z$ .

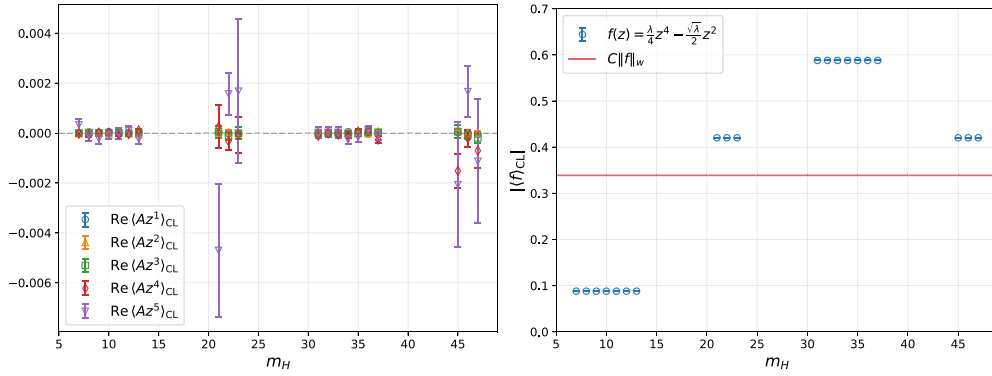
In order to test for correctness with the new criterion, we need to establish the validity of two requirements: the SDEs (24) and the bounds (26). For the former, we consider monomial observables  $z^k$  with integer  $k > 0$ , for which the SDEs read

$$\langle Az^k \rangle = k \langle z^{k-1} \rangle - \lambda \langle z^{k+3} \rangle = 0, \tag{67}$$

with  $A$  given in (12). The left-hand side of this equation, as resulting from complex Langevin simulations (with complex Langevin expectation values labeled by the subscript ‘CL’), is plotted as a function of the kernel parameter  $m_H$  in figure 1(left). Notice that we only show results for those  $m_H$  for which there are no boundary terms [43]. In the cases where boundary terms are nonzero, we found the SDEs to be strongly fluctuating, making it impossible to obtain a usable signal. Also note that fluctuations increase with increasing  $k$  such that we refrain from showing results for  $k > 5$ . Both of these observations, however, are consistent with the expectations.

The main message of figure 1(left) is that there are four different regions in  $m_H$  for which the SDEs are well satisfied and boundary terms vanish. Nonetheless, it is known from [48] that

<sup>6</sup> In the cited works, it was  $K \equiv H^2$  that was parameterized instead of  $H$ , but here all definitions are adapted accordingly to produce an equivalent outcome.



**Figure 1.** Complex Langevin results for the model (64), with  $\lambda = e^{5i\pi/6}$  and a kernel of the form (66), as a function of the parameter  $m_H$ . We only show results for values of  $m_H$  for which boundary terms are consistent with zero. (left): validity of the SDEs (67) with  $k = 1, \dots, 5$ . Note that we only plot the real parts of  $\langle Az^k \rangle$  here; the imaginary parts show similar behavior. The dashed horizontal line marks zero. (right): absolute value of the expectation value of the control observable  $f(z)$  in (72). The solid horizontal line shows the bound  $C||f||_w$  from (26).

only one of these regions corresponds to correct solutions. Thus, we are indeed confronted with a situation where the study of boundary terms (or SDEs) alone is not sufficient to make definite statements about correct convergence. This was found to be the case also in another simple model in [50]. The same is true for the widely-employed correctness criterion based on the decay properties of the complex-Langevin drift term [33]. Indeed, as was shown in [43], the respective distributions of  $z$  in the complex plane for, e.g.  $m_H = 10$  and  $m_H = 34$  are identical up to a rotation. This can be understood from the fact that for both these choices of kernel the complex Langevin evolution is equivalent to a real Langevin evolution, albeit on a rotated real axis, the rotation angle being the only thing distinguishing between them. As a consequence, the distributions of drift magnitudes must be equivalent by definition. Clearly, thus, the drift criterion cannot detect this incorrect convergence. Let us therefore invoke the new correctness criterion.

To this end, we follow section 2.3 by splitting

$$\rho(z) = \frac{1}{Z} e^{-S(z)}, \quad Z = \int_{\mathcal{M}_r} dz e^{-S(z)} \quad (68)$$

according to (18), with the choice

$$w(z) = e^{-\frac{S(z)}{2}}, \quad \rho_r(z) = \frac{1}{Z} e^{-\frac{S(z)}{2}}. \quad (69)$$

With this, the constant  $C$  in (20) is simply given by

$$C = \frac{1}{|Z|} \int_{\mathcal{M}_r} dz \left| e^{-\frac{S(z)}{2}} \right| = \frac{\int_{-\infty}^{\infty} d\xi e^{-\frac{\xi^4}{8}}}{\left| \int_{-\infty}^{\infty} d\xi e^{-\frac{\xi^4}{4}} \right|} = \sqrt[4]{2}, \quad (70)$$

which is independent of  $\lambda$  due to our  $\lambda$ -dependent choice of  $\mathcal{M}_\tau$ . Moreover, the norm  $\|f\|_w$  in (19),

$$\|f\|_w = \sup_{z \in \mathcal{M}_\tau} |f(z)| e^{-\frac{S(z)}{2}}, \quad (71)$$

can straightforwardly be evaluated numerically for polynomial  $f(z)$ . Let us consider the following observable:

$$f(z) = \frac{\lambda}{4} z^4 - \frac{\sqrt{\lambda}}{2} z^2, \quad (72)$$

for which we find  $C\|f\|_w \approx 0.34$  for the bound on the right-hand side of (26). In figure 1(right), we show the  $m_H$ -dependence of the complex Langevin expectation value of  $f$  for those values of  $m_H$  for which the SDEs are valid. Moreover,  $C\|f\|_w$  is shown as the full horizontal line for comparison.

One finds that out of the four regions on which the SDEs hold, the bound (26) is obeyed only on one, namely the one around  $m_H = 10$ . This is precisely the range of  $m_H$  on which the complex Langevin evolution produces correct results [48], i.e., samples only the desired integration cycle and no other [43]. Thus, with this particular choice of observable, the proposed correctness criterion is capable of ruling out the other plateaus. We emphasize once again, however, that the results of figure 1(right) are not sufficient on their own to prove correctness on the plateau around  $m_H = 10$ . To unambiguously prove correctness, one would have to check an infinite number of bounds.

**5.1.2. Two dimensions.** A very similar analysis can be carried out for a straightforward generalization of (64) to two dimensions that was also studied in [43]:

$$S(x_1, x_2) = \frac{\lambda}{4} (x_1^2 + x_2^2)^2. \quad (73)$$

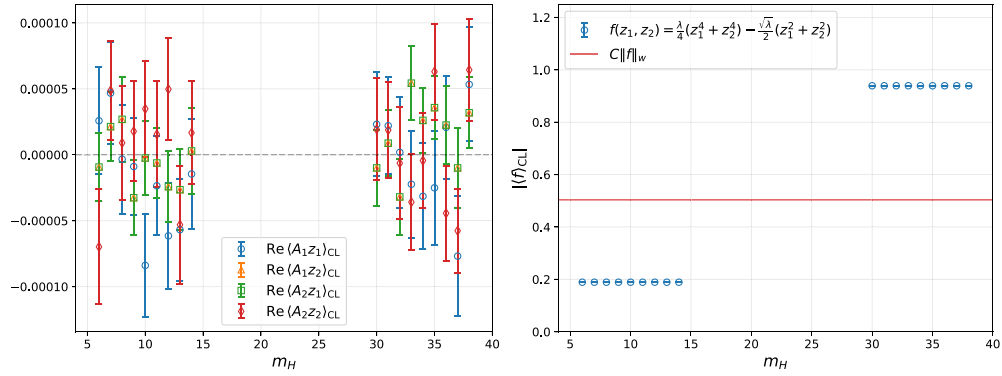
As discussed in [43], there are only two independent integration cycles in this model. For real and positive  $\lambda$ , they correspond to the two manifolds on which, after complexification,  $z_1$  and  $z_2$  are both real and both imaginary, respectively. For general  $\lambda$ , we rotate these cycles similarly to the one-dimensional case discussed above. In particular, the cycle  $\gamma_1 = \mathcal{M}_\tau$  is defined such that on  $\gamma_1$  the action is real. Once again, in the following we use the notation  $z_i$  for the degrees of freedom exclusively.

The Langevin equation we employ reads

$$dz_i(\tau) = -H^2 \frac{\partial S}{\partial z_i} d\tau + Hdw_i(\tau), \quad (74)$$

i.e. we assume the kernel to be proportional to the identity matrix,  $H_{ij} = \delta_{ij}H$ , and we parametrize  $H$  as in (66). For testing the SDEs, we consider monomial observables of the form  $z_1^{k_1} z_2^{k_2}$  with integer  $k_i \geq 0$ . The SDEs (24) then read

$$\begin{aligned} \left\langle A_1 z_1^{k_1} z_2^{k_2} \right\rangle &= k_1 \left\langle z_1^{k_1-1} z_2^{k_2} \right\rangle - \lambda \left[ \left\langle z_1^{k_1+3} z_2^{k_2} \right\rangle + \left\langle z_1^{k_1+1} z_2^{k_2+2} \right\rangle \right] = 0, \\ \left\langle A_2 z_1^{k_1} z_2^{k_2} \right\rangle &= k_2 \left\langle z_1^{k_1} z_2^{k_2-1} \right\rangle - \lambda \left[ \left\langle z_1^{k_1+2} z_2^{k_2+1} \right\rangle + \left\langle z_1^{k_1} z_2^{k_2+3} \right\rangle \right] = 0. \end{aligned} \quad (75)$$



**Figure 2.** Similar to figure 1 but for the model (73). We again use  $\lambda = e^{5i\pi/6}$  and a kernel proportional to the identity matrix, its entries given by (66). (left): validity of (75) with  $k_1 = 0, k_2 = 1$  and vice versa. (right): expectation value of  $f(z)$  in (78) and comparison with  $C||f||_w$  from (26).

The left-hand sides of (75) as a function of the kernel parameter  $m_H$  for the simple cases  $(k_1, k_2) = (0, 1)$  and  $(1, 0)$  are shown in figure 2(left). Again, we find different extended regions in  $m_H$  on which the SDEs hold and boundary terms vanish (for values of  $m_H$  where there are nonzero boundary terms again no results are shown). Here, however, there are only two such regions. This has to do with the way the number of independent integration cycles is reduced from three in one dimension down to two in the two-dimensional model (73) and is discussed in [43].

In order to apply the correctness criterion, we split  $\rho(z_1, z_2)$  into

$$w(z_1, z_2) = e^{-\frac{s(z_1, z_2)}{2}}, \quad \rho_r(z_1, z_2) = \frac{1}{Z} e^{-\frac{s(z_1, z_2)}{2}}, \quad (76)$$

with

$$Z = \int_{\mathcal{M}_r} dz_1 dz_2 e^{-s(z_1, z_2)}, \quad (77)$$

similar to before, and we obtain  $C = \sqrt{2}$  from (20). As a control observable, we choose

$$f(z_1, z_2) = \frac{\lambda}{4} (z_1^4 + z_2^4) - \frac{\sqrt{\lambda}}{2} (z_1^2 + z_2^2), \quad (78)$$

yielding  $C||f||_w \approx 0.50$ . The analog of figure 1(right) for the two-dimensional model is shown in figure 2(right), with equivalent conclusions.

We have thus established that the new correctness criterion can detect incorrect convergence due to unwanted integration cycles in simple quartic models in one and two dimensions. We believe that these results can be extended to quartic models in higher dimensions without too many complications. Regarding models with more involved polynomial actions, we are confident that one should be able to find viable control observables  $f(\mathbf{z})$  reasonably easily, but we have not attempted to do so yet. We stress that  $f(z)$  and  $f(z_1, z_2)$  above were chosen in a purely *ad-hoc* manner and not derived systematically. At the time of writing, we are not aware of any good guiding principle on how to choose appropriate  $f(\mathbf{z})$  for general theories. This will become particularly clear for the model studied in the next subsection.

### 5.2. Densities with zeros

We now turn to a (one-dimensional) theory whose density  $\rho$  has a zero on  $\mathcal{M}_c$ . In particular, we shall be concerned with the model described by

$$\rho(z) = \frac{1}{Z} (z - z_p)^{n_p} e^{-\beta z^2}, \quad Z = \int_{-\infty}^{\infty} dz (z - z_p)^{n_p} e^{-\beta z^2}. \quad (79)$$

This theory, for which again  $\mathcal{M}_r = \mathbb{R}$ , was studied before in [8, 34]. In particular, it was found in [34] that the complex Langevin evolution (without the introduction of a kernel and for  $z_p$  away from the real axis) converges correctly for large enough values of  $\beta$ , but fails for smaller ones. This failure, despite the absence of boundary terms, was then explained to be due to the contribution of unwanted integration cycles [8, 34], making the model a prime target for the new correctness criterion. We emphasize that this is an example for the appearance of unwanted cycles even without using a kernel. Let us now discuss the setup in detail.

The density (79), in which  $n_p$  is assumed to be a positive integer,  $\beta$  is assumed real and  $z_p$  is an arbitrary complex number, has two linearly independent integration cycles, namely

$$\int_{\gamma_+} dz \equiv \int_{z_p}^{\infty} dz, \quad \int_{\gamma_-} dz \equiv \int_{-\infty}^{z_p} dz, \quad (80)$$

where the exact integration paths are again irrelevant due to Cauchy's theorem. The only important property is that the paths start at some point in complex infinity where  $\rho$  vanishes and end in  $z_p$ , the finite zero of  $\rho$ .

The integral over the real axis, which is what we are ultimately after, can be written as a linear combination of the two. In particular, with the definitions

$$Z_{\pm} \equiv \int_{\gamma_{\pm}} dz (z - z_p)^{n_p} e^{-\beta z^2}, \quad (81)$$

such that  $Z = Z_+ + Z_-$ , we can write the desired expectation values as

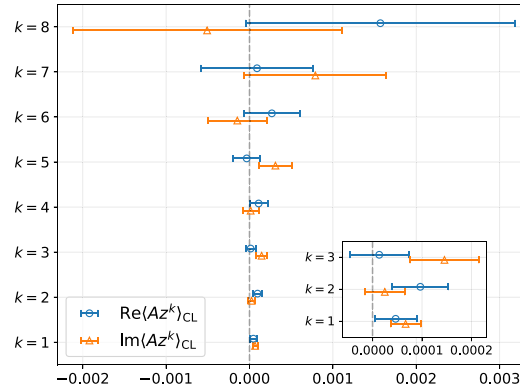
$$\langle f \rangle_{\rho} \equiv \int_{-\infty}^{\infty} dz f(z) \rho(z) = \frac{Z_+}{Z} \langle f \rangle_+ + \frac{Z_-}{Z} \langle f \rangle_-, \quad (82)$$

with

$$\langle f \rangle_{\pm} \equiv \frac{1}{Z_{\pm}} \int_{\gamma_{\pm}} dz f(z) (z - z_p)^{n_p} e^{-\beta z^2}. \quad (83)$$

Therefore, in the light of the theorem (62), the complex Langevin evolution is correct if and only if it produces  $a_{\pm} = \frac{Z_{\pm}}{Z}$ , where we have defined  $\gamma_1 = \gamma_+$ ,  $\gamma_2 = \gamma_-$  and similar for  $a_1$  and  $a_2$  in (62).

In the following, we consider  $\beta = 1.6$ ,  $z_p = i$ , and  $n_p = 2$ , giving  $\frac{Z_{\pm}}{Z} \approx 0.50 \pm 0.23i$ . That these values are not reproduced by a complex Langevin simulation (without a kernel), i.e. that the simulation produces incorrect results, was found in [8]. In the following, we demonstrate that one may use the new correctness criterion to prove this fact without having to compute the  $a_{\pm}$  explicitly. For the simulation of (79), our simulation setup is analogous to the one used in the previous subsection, outlined in [43], the only difference being that here we restrict the maximum allowed Langevin step size in our adaptive step-size algorithm to  $10^{-6}$  instead of  $10^{-5}$ .



**Figure 3.** Validity of the SDEs (84) in the model (79) with  $\beta = 1.6$ ,  $z_p = i$ ,  $n_p = 2$ ,  $k = 1, \dots, 8$  and without a kernel. The inset shows a close-up view of the smallest few  $k$ . The data points have been offset vertically for visual clarity.

In order to apply the correctness criterion, as before, we first study the validity of the SDEs (24), which, for the model (79) and monomial observables  $z^k$ , read

$$\langle Az^k \rangle = k \langle z^{k-1} \rangle + n_p \left\langle \frac{z^k}{z - z_p} \right\rangle - 2\beta \langle z^{k+1} \rangle = 0. \quad (84)$$

This equation demonstrates that here we indeed have to consider a larger space of observables than just the polynomials, since  $\mathcal{H} \neq \mathfrak{A}$  due to the zero of  $\rho(z)$  at  $z = z_p$ . The left-hand side of (84) for different  $k$  is shown in figure 3 and we find that the SDEs are indeed well satisfied.

Now, in order to show that the complex Langevin results are nonetheless incorrect via the new correctness criterion, it again suffices to find a single control observable  $f$  that violates the bounds in (26). We define

$$w(z) = e^{-\frac{\beta z^2}{2}}, \quad \rho_r(z) = \frac{1}{Z} (z - z_p)^{n_p} e^{-\frac{\beta z^2}{2}}. \quad (85)$$

While it turned out that finding such a polynomial can be a nontrivial task in practice, we were able to obtain at least a handful that serve our purpose. The (notably) simplest one of them reads

$$f(z) = \sum_{n=0}^{12} c_n z^n, \quad (86)$$

with the coefficients

$$\begin{aligned} c_0 = 210, \quad c_1 = -1213i, \quad c_2 = -3505, \quad c_3 = 6331i, \quad c_4 = 8583, \\ c_5 = -8175i, \quad c_6 = -6825, \quad c_7 = 3813i, \quad c_8 = 2204, \\ c_9 = -698i, \quad c_{10} = -300, \quad c_{11} = 42.2i, \quad c_{12} = 14.2. \end{aligned} \quad (87)$$

It was obtained by minimizing the distance<sup>7</sup>  $\|f(z) - g(z)\|_w$  to the function  $g(z) = e^{-2(1+(2i+z)^2)}$ , which we found (rather by coincidence) to violate the bounds, within the norm (19). For this polynomial, we find that the complex Langevin evolution produces an expectation value whose absolute value reads  $|\langle f \rangle_{\text{CL}}| = 902$  with negligible statistical uncertainties, while the bound on the right-hand side of (26) reads  $C\|f\|_w \approx 702$ . This bound is thus clearly violated, from which one can safely conclude that the complex Langevin results are incorrect.

## 6. Towards more realistic models

In this section, we briefly comment on the prospects of applying the new correctness criterion to lattice models. After splitting the measure of the theory as in (18), we need to establish whether the following bound holds in a complex Langevin simulation for all observables  $f \in \mathcal{H}$ :

$$|\langle f \rangle_{\text{CL}}| \leq C\|f\|_w, \quad (88)$$

with  $\|f\|_w$  and  $C$  defined in (19) and (20), respectively.

For a given  $f$ , determining  $\|f\|_w$  is relatively easy. For instance, one could start at a random point in  $\mathcal{M}_r$  and find a local maximum of  $|f(\mathbf{x}, \mathbf{g})|w(\mathbf{x})$  using the gradient *ascent* method, since both  $f$  and  $\omega$  are expected to be smooth functions in practice. To find a global maximum, one starts the gradient ascent method at  $N_{\text{init}}$  initial points and chooses the largest local maximum as an approximation to the global maximum. Since  $|f(\mathbf{x}, \mathbf{g})|w(\mathbf{x}) \rightarrow 0$  as  $|\mathbf{x}| \rightarrow \infty$ , this procedure is expected to converge.

In order to compute  $C$ , let us introduce the non-normalized density  $\rho'$  as

$$\rho(\mathbf{x}, \mathbf{g}) \equiv \frac{1}{Z} \rho'(\mathbf{x}, \mathbf{g}) \quad \text{with} \quad Z \equiv \int_{\mathcal{M}_r} \rho'(\mathbf{x}, \mathbf{g}) d\mu, \quad (89)$$

and similar for the reduced density  $\rho_r$ ,

$$\rho_r(\mathbf{x}, \mathbf{g}) \equiv \frac{1}{Z_r} \rho'_r(\mathbf{x}, \mathbf{g}), \quad (90)$$

such that  $\rho'(\mathbf{x}, \mathbf{g}) = w(\mathbf{x})\rho'_r(\mathbf{x}, \mathbf{g})$ . Then, we find that

$$C = \int_{\mathcal{M}_r} |\rho_r(\mathbf{x}, \mathbf{g})| d\mu = \frac{\int_{\mathcal{M}_r} |\rho'_r(\mathbf{x}, \mathbf{g})| d\mu}{\left| \int_{\mathcal{M}_r} \rho'(\mathbf{x}, \mathbf{g}) d\mu \right|} \quad (91)$$

and, thus,

$$C = \frac{\int_{\mathcal{M}_r} |\rho'_r(\mathbf{x}, \mathbf{g})| d\mu}{\left| \int_{\mathcal{M}_r} \frac{\rho'(\mathbf{x}, \mathbf{g})}{|\rho'_r(\mathbf{x}, \mathbf{g})|} |\rho'_r(\mathbf{x}, \mathbf{g})| d\mu \right|} = \frac{1}{\left| \left\langle \frac{\rho'(\mathbf{x}, \mathbf{g})}{|\rho'_r(\mathbf{x}, \mathbf{g})|} \right\rangle_{|\rho'_r|} \right|}. \quad (92)$$

<sup>7</sup> One possible approach to find suitable polynomials in practice could thus be to start with an arbitrary function for which one can somehow show that it violates the bounds (26) and then approximate it (in the norm (19)) with a polynomial of sufficiently high order. We also remark that we are, in fact, not restricted to using polynomials here. Any element of  $\mathcal{H}$  would work in principle.

Here, we have defined the expectation value with respect to  $|\rho'_r|$  as

$$\langle f(\mathbf{x}, \mathbf{g}) \rangle_{|\rho'_r|} \equiv \frac{\int_{\mathcal{M}_r} f(\mathbf{x}, \mathbf{g}) |\rho'_r(\mathbf{x}, \mathbf{g})| d\mu}{\int_{\mathcal{M}_r} |\rho'_r(\mathbf{x}, \mathbf{g})| d\mu}. \quad (93)$$

This means that  $C$  can in principle be evaluated in a Monte-Carlo simulation using the measure  $|\rho'_r(\mathbf{x}, \mathbf{g})|$  on the manifold  $\mathcal{M}_r$ . Note, however, that  $C$  is the inverse of a quantity that resembles the sign average in a phase-quenched simulation, with the difference that here we consider  $|\rho_r|$  instead of  $|\rho|$ .  $C$  might thus become large in the case of lattice models with a strong sign problem. We mention, however, that one may explicitly introduce the phase  $\varphi'_r$  of  $\rho'_r = |\rho'_r| e^{i\varphi'_r}$  to write

$$\frac{1}{\left\langle \frac{\rho'_r(\mathbf{x}, \mathbf{g})}{|\rho'_r(\mathbf{x}, \mathbf{g})|} \right\rangle_{|\rho'_r|}} = \frac{1}{\langle e^{i\varphi'_r(\mathbf{x}, \mathbf{g})} w(\mathbf{x}) \rangle_{|\rho'_r|}}. \quad (94)$$

The appearance of  $w$  as a damping factor might be beneficial in practical computations.

## 7. Summary and conclusions

In this work, we have derived a criterion for correctness of probability measures  $dP$  on complex manifolds  $\mathcal{M}_c$ , given a complex density  $\rho$  on a real manifold  $\mathcal{M}_r$ . In particular, we have found necessary and sufficient conditions that ensure the correctness of  $dP$ , i.e.

$$\int_{\mathcal{M}_c} dP f(\mathbf{z}, \mathbf{g}_c) = \int_{\mathcal{M}_r} d\mu \rho(\mathbf{x}, \mathbf{g}) f(\mathbf{x}, \mathbf{g}), \quad (95)$$

for holomorphic  $f$  within a certain algebra of observables  $\mathcal{H}$ , defined in (14). Here, we assume  $\mathcal{M}_r$  to be of the product form (1), with  $d\mu$  denoting a suitable integration measure, and  $\mathcal{M}_c$  to be its complexification.

More precisely, the necessary and sufficient conditions are on the one hand the validity of the SDEs (24) and (25), and on the other hand the validity of the bounds (26) for all  $f \in \mathcal{H}$ , which need to hold simultaneously. These results were derived using standard theorems of functional analysis, assuming only that  $\rho$  is holomorphic and normalizable and that the partial derivatives of  $\log \rho$  (with respect to both  $\mathbf{x}$  and  $\mathbf{g}$ ) grow at most like a polynomial on  $\mathcal{M}_r$ . None of these assumptions is particularly restrictive in the view of usual theories of interest in physics. Moreover, we have required  $\mathcal{M}_r$  to be free of any zeros of  $\rho$ , which, however, can always be achieved by a continuous deformation of  $\mathcal{M}_r$ . In particular, the presence of a fermionic determinant in the density, as it appears in the path integral treatment of quantum field theories with fermions, does not spoil the proof. Finally, the quantities  $w$  and  $\rho_r$ , defined in (18), are to be chosen such that the former obeys (21) and that the constant  $C$  in (20) as well as the norms  $\|f\|_w$  in (19) (for all  $f \in \mathcal{H}$ ) are finite.

We should mention that there is a lot of freedom in the definition of our criteria: On the one hand, there is freedom in the choices of  $w$ ,  $\rho_r$ , and  $\mathcal{M}_r$ , each different choice giving rise to different bounds in general. Since different manifolds  $\mathcal{M}_r$  that are equivalent via Cauchy's theorem produce different bounds, one may even attempt to take the minimum of (26) over all equivalent  $\mathcal{M}_r$  to find the lowest bound for given  $w$ ,  $\rho_r$ , and  $f$ . On the other hand, one may define  $\|f\|_w$  and  $C$  using different norms than those in (19) and (20), see, e.g. (54) and (55). Any violation of the bounds for any valid choice of  $\mathcal{M}_r$ ,  $w$ ,  $\rho_r$ ,  $f$ , as well as of the norms implies that (95) cannot be true.

We have then proceeded to test our chosen criterion in the context of complex Langevin simulations. The complex Langevin expectation values are naturally defined via a probability measure  $dP$ , obtained from the stationary solution of the real Fokker–Planck equation, hence making them amenable to such an investigation. Indeed, we have found for simple one- and two-dimensional toy models, for which exact results are readily available for comparison, that the criterion is capable of detecting incorrect convergence where conventional correctness criteria, such as boundary-term analyses, fail. In particular, this is the case when the complex Langevin expectation values satisfy the SDEs but receive contributions from unwanted integration cycles  $\gamma_i \neq \mathcal{M}_r$ . To the best of our knowledge, there are no other correctness criteria for complex Langevin available that would be sensitive to different integration cycles. Whether unwanted integration cycles play a role in realistic lattice models or even gauge theories remains an open question. A few exploratory results in this direction can be found in [43, 51].

Finally, we emphasize once again that our criterion is not suitable for proving that a given simulation produces correct results, as this would require checking an infinite amount of bounds (26). What we have demonstrated, however, is that it is capable of ruling out correctness for suitable choices of  $f$  without the need of computing exact results for observables. How to find such  $f$  for a given theory is a nontrivial question, to which we cannot yet provide a definite answer.

### Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: [10.5281/zenodo.16874951](https://doi.org/10.5281/zenodo.16874951); [10.5281/zenodo.16874895](https://doi.org/10.5281/zenodo.16874895). Our simulation code is available upon request.

### Acknowledgment

The authors thank Lorenzo Salcedo for useful comments on the manuscript. This research was funded in whole, or in part, by the Austrian Science Fund (FWF) [10.55776/P36875]. The numerical simulations, from which the results discussed in this work have been computed, were performed on the computing cluster of the University of Graz (GSC). The data analyses are based on the Python scientific computing ecosystem [52], in particular via the packages [53–55] and we are also grateful for the creation and maintenance of the dependencies of the latter. Moreover, auxiliary numerical calculations were performed with the Wolfram Mathematica computer algebra system [56].

### Open access statement

For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

### Appendix A. Proof of continuity of $w$

Here, we prove the continuity of  $w$  as defined in (22).

**Proof.** It suffices to show continuity of  $\rho_m(\mathbf{x}) = \sup_{\mathbf{g} \in G^m} |\rho(\mathbf{x}, \mathbf{g})|$ . Because  $G^m$  is compact and  $\rho$  is continuous, the supremum is attained at some  $\mathbf{g}_x \in G^m$  for every  $\mathbf{x}$ , i.e.

$$\rho_m(\mathbf{x}) = |\rho(\mathbf{x}, \mathbf{g}_x)|. \quad (\text{A.1})$$

We have to show that for any  $\mathbf{x} \in \mathbb{R}^n$  and any  $\epsilon > 0$  there is a  $\delta(\epsilon) > 0$  such that

$$|\rho_m(\mathbf{x}) - \rho_m(\mathbf{y})| < \epsilon \quad \forall \mathbf{y} \quad \text{with} \quad |\mathbf{x} - \mathbf{y}| < \delta(\epsilon). \quad (\text{A.2})$$

According to lemma 1 below,  $|\rho(\mathbf{x}, \mathbf{g})|$  is continuous in  $\mathbf{x}$ , uniformly in  $\mathbf{g} \in G^m$ , i.e. for any  $\epsilon > 0$  there is a  $\delta(\epsilon) > 0$  such that

$$|\rho(\mathbf{x}, \mathbf{g}_x) - \rho(\mathbf{y}, \mathbf{g}_x)| < \epsilon \quad \forall \mathbf{y} \quad \text{with} \quad |\mathbf{x} - \mathbf{y}| < \delta(\epsilon) \quad (\text{A.3})$$

and

$$|\rho(\mathbf{x}, \mathbf{g}_y) - \rho(\mathbf{y}, \mathbf{g}_y)| < \epsilon \quad \forall \mathbf{y} \quad \text{with} \quad |\mathbf{x} - \mathbf{y}| < \delta(\epsilon). \quad (\text{A.4})$$

Thus

$$\begin{aligned} \rho_m(\mathbf{x}) &= |\rho(\mathbf{x}, \mathbf{g}_x)| < |\rho(\mathbf{y}, \mathbf{g}_x)| + \epsilon \leq \rho_m(\mathbf{y}) + \epsilon \\ &< |\rho(\mathbf{x}, \mathbf{g}_y)| + 2\epsilon \leq \rho_m(\mathbf{x}) + 2\epsilon, \end{aligned} \quad (\text{A.5})$$

provided  $|\mathbf{x} - \mathbf{y}| < \delta(\epsilon)$ . This implies

$$\rho_m(\mathbf{x}) - \epsilon < \rho_m(\mathbf{y}) < \rho_m(\mathbf{x}) + \epsilon, \quad (\text{A.6})$$

which shows the continuity of  $w$  defined by (22).  $\square$

In the proof we used the

**Lemma 1.**  $\rho(\mathbf{x}, \mathbf{g})$  is continuous in  $\mathbf{x}$ , uniformly in  $\mathbf{g}$ . This means that for any  $\epsilon > 0$  there is a  $\delta(\epsilon) > 0$ , independent of  $\mathbf{g}$ , such that

$$|\rho(\mathbf{x}, \mathbf{g}) - \rho(\mathbf{y}, \mathbf{g})| < \epsilon \quad \forall \mathbf{y} \quad \text{with} \quad |\mathbf{x} - \mathbf{y}| < \delta(\epsilon). \quad (\text{A.7})$$

**Proof.** Since we only need uniformity in  $\mathbf{g}$ , we use the fact that  $\mathbf{x}$  always lies in the interior of a compact set  $B \subset \mathbb{R}^n$ . Then the lemma follows from the standard theorem that a continuous function on a compact set is uniformly continuous; see, e.g. [57] (theorem 4.19). The fact that  $\delta(\epsilon)$  may depend on  $B$  is of no consequence.  $\square$

## Appendix B. Approximation in the $w$ norm

In this appendix, we demonstrate that the functions  $f_{\mathbf{k}}$  defined in (44) are contained in the closure  $\bar{\mathfrak{A}}$  of  $\mathfrak{A}$  with respect to the norm  $\|\cdot\|_w$ .

**Proof.** To see this, we have to show that  $\exp(i\mathbf{k} \cdot \mathbf{x})$  is a limit of polynomials, which follows by considering the Taylor expansion

$$\exp(i\mathbf{k} \cdot \mathbf{x}) = \sum_{n=0}^N \frac{(i\mathbf{k} \cdot \mathbf{x})^n}{n!} + R_N \quad (\text{B.8})$$

with  $t \in \mathbb{R}$  and a remainder  $R_N$ . By using (21), one can show that  $\|R_N\|_w$  satisfies

$$\|R_N\|_w \leq \frac{B}{(N+1)!} e^{-\frac{N+1}{2}} \left(\frac{N+1}{2b}\right)^{\frac{N+1}{2}} |\mathbf{k}|^{N+1}. \tag{B.9}$$

Here,  $B$  and  $b$  are the constants defined in (21) and we have used the general bound on the remainder in Taylor’s theorem for a function  $f(t)$ , given by

$$|R_N| \leq M \frac{|t^{N+1}|}{(N+1)!}, \tag{B.10}$$

provided

$$|f^{(N+1)}(\tau)| \leq M \quad \text{for } |\tau| \leq |t|. \tag{B.11}$$

In our case, we have  $M = |\mathbf{x}|^{N+1} |\mathbf{k}|^{N+1}$  and  $t = 1$ . Since  $\lim_{N \rightarrow \infty} \|R_N\|_w = 0$ , this shows that we obtain an arbitrarily good polynomial approximation in the  $\|\cdot\|_w$  norm for functions of the form (44).  $\square$

### Appendix C. Boundary terms, convergence conditions, and Schwinger–Dyson equations

Here, we collect a number of facts regarding the relation between vanishing boundary terms, the validity of the SDEs and the validity of the so-called convergence conditions (CCs) [28]. The latter, also known as consistency conditions, are an expression of the fact that the complex Langevin evolution has reached equilibrium. More precisely, we introduce the following

**Definition.** A linear functional  $T_{CC}$  on  $\mathcal{H}$  is said to satisfy the CCs if

$$(T_{CC}, L_c f) = 0 \quad \forall f \in \mathcal{H}, \tag{C.12}$$

where

$$L_c \equiv \sum_{i=1}^n A_i \partial_i + \sum_{j=1}^d \sum_{k=1}^m A_{j,k} D_{j,k} \tag{C.13}$$

is the so-called complex Langevin operator (in the second term it was assumed that the group derivations are with respect to a basis of the Lie algebra which is orthogonal with respect to the Killing form) and the  $A_i$  and  $A_{j,k}$  are the SD operators defined in (12) and (13), respectively.

If there are no boundary terms, the CCs (C.12) follow from the statement that the measure  $dP$  on  $\mathcal{M}_c$  satisfies the time-independent Fokker–Planck equation. The converse, however, is not necessarily true, as there are examples in which the CCs hold for some observables even though they come with nonzero boundary terms.

If the SDEs are satisfied, the CCs follow automatically, as is trivially shown [28]. Moreover, in [28] it was also argued that the converse is true as well. Unfortunately, the argument given there is incorrect. Thus, in the following, we derive a sufficient condition under which the CCs do imply the SDEs, using the methods employed in the main text.

**Proposition.** Let  $dP$  be a probability measure on  $\mathcal{M}_c$ . Assume that  $\rho$  does not vanish on  $\mathcal{M}_r$  and that the derivatives of  $\log \rho$  grow at most polynomially in  $|\mathbf{x}|$ , that  $dP$  obeys the bounds (26), and that the linear functional on  $\mathcal{H}$  generated by  $dP$  satisfies the CCs (C.12). Then, for all  $f \in \mathcal{H}$ ,

$$\int_{\mathcal{M}_c} f dP = \int_{\mathcal{M}_r} f(\mathbf{x}) w(\mathbf{x}) \tilde{\sigma}_{\text{CC}}(\mathbf{x}, \mathbf{g}) d\mu, \quad (\text{C.14})$$

where  $\tilde{\sigma}_{\text{CC}}$  is a complex measure on  $\mathcal{M}_r$  of finite total variation. Furthermore, defining

$$\sigma_{\text{CC}}(\mathbf{x}, \mathbf{g}) \equiv w(\mathbf{x}) \tilde{\sigma}_{\text{CC}}(\mathbf{x}, \mathbf{g}), \quad (\text{C.15})$$

$\sigma_{\text{CC}}$  satisfies

$$L_c^T \sigma_{\text{CC}} = 0, \quad (\text{C.16})$$

where  $L_c^T$  (the complex Fokker–Planck operator) is the transpose of  $L_c$ ,

$$L_c^T = - \sum_{i=1}^n \partial_i A_i^T - \sum_{j=1}^d \sum_{k=1}^m D_{j,k} A_{j,k}^T. \quad (\text{C.17})$$

**Proof.** The proof is a straightforward imitation of the proofs of propositions 1 and 2 in the main text.  $\square$

As with the SDEs, (C.16) is solved by

$$\sigma_{\text{CC}}(\mathbf{x}, \mathbf{g}) = c \rho(\mathbf{x}, \mathbf{g}), \quad c \in \mathbb{C}, \quad (\text{C.18})$$

where  $c = 1$  due to the normalization of  $P$ . Crucially, this solution also obeys the SDEs. If the eigenvalue 0 of  $L_c^T$  is nondegenerate, the SDEs are therefore implied by the CCs. This means that we can formulate the following

**Corollary.** If a bounded linear functional  $T_{\text{CC}}$  on  $\mathcal{H}$ , generated by a probability density  $dP$  that satisfies the bounds (26), obeys the CCs (C.12), and if the equation (C.16) has only the solution (C.18), then  $T_{\text{CC}}$  also obeys the SDEs.

**Remark.** While we cannot rule out that the solution (C.18) is not unique, this would mean degeneracy of the zero eigenvalue of  $L_c^T$  restricted to the real integration manifold  $\mathcal{M}_r$ . This could not be the case generically; it might at worst happen for exceptional parameter values. The only cases in which such a degeneracy is known to occur, are those in which  $\rho$  vanishes somewhere on  $\mathcal{M}_r$ , leading to nonergodicity; but we have excluded this possibility.

To summarize, on the one hand, if the complex Langevin process relaxes to an equilibrium distribution for which there are no boundary terms, the CCs hold. If in addition the bounds (26) are satisfied and the zero eigenvalue of  $L_c^T$  is nondegenerate, the SDEs hold. On the other hand, the validity of the SDEs always implies the CCs. For all we know, neither the CCs nor the SDEs can guarantee the absence of boundary terms, nor does the absence of boundary terms automatically imply the SDEs. However, at the time of writing, we have never encountered any counterexamples to either of these scenarios.

## ORCID iDs

Michael Mandl  0000-0003-4290-4246

Erhard Seiler  0000-0002-4478-6397

Dénes Sexty  0000-0002-0521-816X

## References

- [1] Parisi G 1983 On complex probabilities *Phys. Lett. B* **131** 393
- [2] Klauder J R 1983 A Langevin approach to fermion and quantum spin correlation functions *J. Phys. A: Math. Gen* **16** L317
- [3] Salcedo L L 1997 Representation of complex probabilities *J. Math. Phys.* **38** 1710
- [4] Weingarten D 2002 Complex probabilities on  $R^N$  as real probabilities on  $C^N$  and an application to path integrals *Phys. Rev. Lett.* **89** 240201
- [5] Salcedo L L 2007 The existence of positive representations for complex weights *J. Phys. A: Math. Theor.* **40** 9399
- [6] Wosiek J 2016 Beyond complex Langevin equations: from simple examples to positive representation of Feynman path integrals directly in the Minkowski time *J. High Energy Phys.* [JHEP04\(2016\)146](#)
- [7] Seiler E and Wosiek J 2017 Positive representations of a class of complex measures *J. Phys. A: Math. Theor.* **50** 495403
- [8] Salcedo L L and Seiler E 2019 Schwinger–Dyson equations and line integrals *J. Phys. A: Math. Theor.* **52** 035201
- [9] Folland G B 1999 *Real Analysis, Modern Techniques and Their Applications* (Wiley)
- [10] Rudin W 1991 *Functional Analysis* (MacGraw-Hill)
- [11] Rudin W 1987 *Real and Complex Analysis* (MacGraw-Hill)
- [12] Reed M and Simon B 1972 *Methods of Mathematical Physics* vol I (Academic)
- [13] Bump D 2013 *Lie Groups* (Springer)
- [14] Knapp A W 1986 *Representation Theory of Semisimple Groups* (Princeton University Press)
- [15] Troyer M and Wiese U-J 2005 Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations *Phys. Rev. Lett.* **94** 170201
- [16] de Forcrand P 2010 Simulating QCD at finite density *PoS (LAT2009)* 010
- [17] Gattringer C and Langfeld K 2016 Approaches to the sign problem in lattice field theory *Int. J. Mod. Phys. A* **31** 1643007
- [18] Berger C E, Rammelmüller L, Loheac A C, Ehmman F, Braun J and Drut J E 2021 Complex Langevin and other approaches to the sign problem in quantum many-body physics *Phys. Rep.* **892** 1
- [19] Parisi G and Wu Y-S 1981 Perturbation theory without gauge fixing *Sci. Sin.* **24** 483
- [20] Batrouni G G, Katz G R, Kronfeld A S, Lepage G P, Svetitsky B and Wilson K G 1985 Langevin simulations of lattice field theories *Phys. Rev. D* **32** 2736
- [21] Salcedo L L 1993 Spurious solutions of the complex Langevin equation *Phys. Lett. B* **305** 125
- [22] Gausterer H and Lee S 1993 The mechanism of complex Langevin simulations *J. Stat. Phys.* **73** 147
- [23] Fujimura K, Okano K, Schülke L, Yamagishi K and Zheng B 1994 On the segregation phenomenon in complex Langevin simulation *Nucl. Phys. B* **424** 675
- [24] Gausterer H 1998 Complex Langevin: a numerical method? *Nucl. Phys. A* **642** c239
- [25] Aarts G, Seiler E and Stamatescu I-O 2010 Complex Langevin method: when can it be trusted? *Phys. Rev. D* **81** 054508
- [26] Aarts G, James F A, Seiler E and Stamatescu I-O 2010 Adaptive stepsize and instabilities in complex Langevin dynamics *Phys. Lett. B* **687** 154
- [27] Aarts G and James F A 2010 On the convergence of complex Langevin dynamics: the three-dimensional XY model at finite chemical potential *J. High Energy Phys.* [JHEP08\(2010\)020](#)
- [28] Aarts G, James F A, Seiler E and Stamatescu I-O 2011 Complex Langevin: etiology and diagnostics of its main problem *Eur. Phys. J. C* **71** 1756
- [29] Aarts G, James F A, Pawłowski J M, Seiler E, Sexty D and Stamatescu I-O 2013 Stability of complex Langevin dynamics in effective models *J. High Energy Phys.* [JHEP03\(2013\)073](#)
- [30] Aarts G, Giudice P and Seiler E 2013 Localised distributions and criteria for correctness in complex Langevin dynamics *Ann. Phys.* **337** 238

- [31] Nishimura J and Shimasaki S 2015 New insights into the problem with a singular drift term in the complex Langevin method *Phys. Rev. D* **92** 011501(R)
- [32] Salcedo L L 2016 Does the complex Langevin method give unbiased results? *Phys. Rev. D* **94** 114505
- [33] Nagata K, Nishimura J and Shimasaki S 2016 Argument for justification of the complex Langevin method and the condition for correct convergence *Phys. Rev. D* **94** 114515
- [34] Aarts G, Seiler E, Sexty D and Stamatescu I-O 2017 Complex Langevin dynamics and zeroes of the fermion determinant *J. High Energy Phys.* **JHEP05(2017)044**
- [35] Scherzer M, Seiler E, Sexty D and Stamatescu I-O 2019 Complex Langevin and boundary terms *Phys. Rev. D* **99** 014512
- [36] Scherzer M, Seiler E, Sexty D and Stamatescu I-O 2020 Controlling complex Langevin simulations of lattice models by boundary term analysis *Phys. Rev. D* **101** 014501
- [37] Seiler E 2020 Complex Langevin: boundary terms at poles *Phys. Rev. D* **102** 094507
- [38] Seiler E, Sexty D and Stamatescu I-O 2024 Complex Langevin: correctness criteria, boundary terms, and spectrum *Phys. Rev. D* **109** 014509
- [39] Boguslavski K, Hotzy P and Müller D I 2025 Lefschetz thimble-inspired weight regularizations for complex Langevin simulations *SciPost Phys.* **18** 092
- [40] Alvestad D, Larsen R and Rothkopf A 2021 Stable solvers for real-time complex Langevin *J. High Energy Phys.* **JHEP08(2021)138**
- [41] Seiler E, Sexty D and Stamatescu I-O 2013 Gauge cooling in complex Langevin for lattice QCD with heavy quarks *Phys. Lett. B* **723** 213
- [42] Flower J, Otto S W and Callahan S 1986 Complex Langevin equations and lattice gauge theory *Phys. Rev. D* **34** 598
- [43] Hansen M W, Mandl M, Seiler E and Sexty D 2025 Role of integration cycles in complex Langevin simulations *Phys. Rev. D* **111** 074502
- [44] Witten E 2011 Analytic continuation Of Chern-Simons theory, *AMS/IP Stud. Adv. Math.* **50** 347
- [45] Wilkinson M D et al 2016 The FAIR guiding principles for scientific data management and stewardship *Sci. Data* **3** 160018
- [46] Mandl M, Seiler E and Sexty D 2025 Dataset for “necessary and sufficient conditions for correctness of complex Langevin” *Zenodo* (<https://doi.org/10.5281/zenodo.16874951>)
- [47] Mandl M, Seiler E and Sexty D 2025 Analysis and visualization code for “necessary and sufficient conditions for correctness of complex Langevin” (<https://doi.org/10.5281/zenodo.16874895>)
- [48] Okamoto H, Okano K, Schülke L and Tanaka S 1989 The role of a kernel in complex Langevin systems *Nucl. Phys. B* **324** 684
- [49] Mandl M, Hansen M W, Seiler E and Sexty D 2024 Kernels and integration cycles in complex Langevin simulations *PoS (LATTICE2024)* 059
- [50] Alvestad D, Larsen R and Rothkopf A 2023 Towards learning optimized kernels for complex Langevin *J. High Energy Phys.* **JHEP04(2023)057**
- [51] Guralnik G and Pehlevan C 2009 Complex Langevin equations and Schwinger–Dyson equations *Nucl. Phys. B* **811** 519
- [52] van Rossum G and de Boer J 1991 Interactively testing remote servers using the python programming language *CWI Q.* **4** 283
- [53] van der Walt S, Colbert S C and Varoquaux G 2011 The numpy array: a structure for efficient numerical computation *Comput. Sci. Eng.* **13** 22
- [54] McKinney W 2010 Data structures for statistical computing in Python *Proc. 9th Python in Science Conf. (Scipy 2010)* p 51
- [55] Hunter J D 2007 Matplotlib: a 2D graphics environment *Comput. Sci. Eng.* **9** 90
- [56] Wolfram Research, Inc. Mathematica, Version 14.2
- [57] Rudin W 1976 *Principles of Mathematical Analysis* (MacGraw-Hill)