PHYSICAL REVIEW D **89,** 105003 (2014)

Perturbative roughness corrections to electromagnetic Casimir energies

Hua Yao Wu and Martin Schaden

Department of Physics, Rutgers University, 101 Warren Street, Newark, New Jersey 07102, USA

(Received 21 February 2014; published 6 May 2014)

Perturbative corrections to the Casimir free energy due to macroscopic roughness of dielectric interfaces are obtained in the framework of an effective low-energy field theory. It describes the interaction of electromagnetic fields with materials whose plasma frequency ω_p determines the low-energy scale. The naïve perturbative expansion of the single-interface scattering matrix in the variance of the profile is sensitive to short-wavelength components of the roughness correlation function. We introduce generalized counterterms that subtract and correct these high-momentum contributions to the loop expansion. To leading order, the counterterms are determined by the phenomenological plasmon model. The latter is found to be consistent with the low-energy description. The proximity force approximation is recovered in the limit of long correlation length and gives the upper limit for the roughness correction to the Casimir force. The renormalized low-energy theory is insensitive to the high-momentum behavior of the roughness correlation function. Predictions at zero temperature of the improved theory are compared with those of the unrenormalized model and with experiment. The Casimir interaction of interfaces with low levels of roughness is found to be well reproduced by that of flat parallel plates with the *measured* reflection coefficients at a distance that is slightly less than the mean separation of the rough surfaces.

DOI: 10.1103/PhysRevD.89.105003 PACS numbers: 03.70.+k, 42.50.-p, 68.35.Ct, 73.20.Mf

I. INTRODUCTION

Casimir originally [1] obtained the force due to electromagnetic zero-point fluctuations between two large ideal parallel metallic flat surfaces at vanishing temperature. His approach was soon generalized to dielectric surfaces [2,3], finite temperature [3,4], and experimentally more accessible geometries [5]. Because it was unimportant in early Casimir experiments [6], the influence of surface roughness was investigated only later [7,8]. Once Casimir forces were accurately measured with atomic force microscope techniques [9] at plate separations of a few hundred nanometers, this correction could no longer be ignored. Effects due to surface roughness are even more important at the small separations and higher accuracy of recent experiments [10,11]. Increasing experimental [12,13] and theoretical [13–23] effort has since been devoted to understanding this correction. The only rigorous nonperturbative approach to roughness currently is the Proximity Force Approximation (PFA) (and some recent modifications thereof [13,20]). This approximation is accurate when the correlation length l_c of the profile greatly exceeds the average plate separation a as well as the inverse plasma frequency $1/\omega_p$ of the material [14,20]. Most rigorous approaches consider perturbative corrections to the Green's function in powers of σ/a or a derivative expansion of the roughness profile [21]. The limit of very rough plates with $a \gg l_c$ was first considered in Ref. [7] using methods of stochastic calculus.

For stochastic roughness, all perturbative calculations to date [7,8,13,17,18,23] show an increase in the magnitude of the Casimir energy and force with decreasing correlation length l_c. They approach the PFA for $l_c \gg a$ [20] as a lower bound. This behavior corresponds to the dashed curves for the ratio of the roughness correction to the Casimir energy of flat plates shown in Fig. 10. As we have argued in Ref. [22] for a scalar field, such a strengthening of the Casimir force due to roughness is not just counterintuitive but also unphysical. From the point of view of the multiple-scattering expansion, decreasing the correlation length decreases the magnitude of the Casimir force, since the reflection coefficient for backscattering is reduced. Corrections to the free energy of higher order in the loop expansion for scalar fields are of similar magnitude [22,23] at small l_c. Assuming the scalar model to be valid at any scale, the leading contributions in $\sigma^2/(l_c a)$ were resummed [22]. In the Casimir energy, this effectively amounted to reducing the separation between two flat plates. The resummed Casimir energy in this approximation indeed decreases with increasing roughness [22]. We will find a similar behavior for the roughness correction to the electromagnetic Casimir force with the present approach.

The perturbative analysis for electromagnetic fields in Refs. [7,8,18] also predicts a strengthening of the Casimir force with increasing roughness. This trend does not appear to be supported by experiment [10,12]. Most experimental investigations [24,25] using machined unidirectional surfaces are in a nonperturbative regime. However, a perturbative analysis of roughness can be justified for some investigations [11,12] that use relatively thin rough gold coatings. The roughness correction compared to flat plates observed at small separations in this case is of the order of 30% only. These experiments appear to measure a Casimir force that is smaller (not larger) than the PFA estimate.

We here set out to extend the low-energy formalism we developed for scalar fields to interactions of the electromagnetic field with matter. Apart from being more complicated, the basic field-theoretic approach is similar. However, there is a fundamental difference between the effective low-energy theories for scalar and electromagnetic fields: whereas one can pretend that the scalar theory is valid at all length scales, this is not possible for the electromagnetic model. The coupling to the roughness profile sets the energy scale in the scalar model, whereas this interaction for the electromagnetic case is dimensionless. The scale of the low-energy effective electromagnetic theory is the plasma frequency, ω_p. The description of electromagnetic interactions with materials by their dielectric permittivity is not reasonable for momentum and energy transfers much above ω_p. Explicitly resumming high-momentum contributions to the loop expansion thus would be quite out of control in the low-energy effective electromagnetic model.

Unfortunately, the roughness contribution to the low-energy electromagnetic scattering matrix arises to a large part from (loop) momenta $q \sim 1/l_c \gtrsim \omega_p$. Using the low-energy theory to compute these high-momentum contributions is not justified. Instead of resumming high orders of the loop expansion, we use generalized counterterms to correct for the high-momentum contributions [26–28]. One thus trades (wrong) high-momentum contributions to the scattering matrix for a phenomenological description—in this case, plasmon scattering.

The observation that UV-divergent vacuum energies arise due to unphysical boundary conditions is quite old [29]. Some UV divergences may be absorbed in the renormalization of physical parameters [30]. Sometimes they are avoided by a more realistic modeling of the surface. However, this invariably gives results that are *sensitive* to the modeling of the interactions with materials at high energies. Here we address the related, but somewhat different, issue that the low-energy description is not suited for computing high-momentum contributions to physical observables, whether they diverge or not. High-momentum parts of loop integrals should not be viewed as reliable predictions of an effective low-energy theory and generally have to be corrected phenomenologically.

This article is organized as follows: In Sec. II, we present Schwinger's low-energy theory for electromagnetic interactions with materials and derive the scattering matrix $\mathbf{T}^h$ for roughness corrections to the Casimir free energy. The one-loop correction to $\mathbf{T}^h$ is found to be UV sensitive in Sec. III. We show that this problem may be solved by subtraction and inclusion of a phenomenological plasmon contribution. In Sec. IV, roughness corrections to the Casimir free energy to the leading order of the variance are derived. They differ from earlier results by the inclusion of a counterterm that corrects uncontrolled high-momentum contributions to loop integrals. We obtain the limits of very large and very small correlation lengths as well as the ideal metal limit, and we determine the plasmon coupling at low energies by analyticity arguments. Section V develops the low-energy effective field theory of electromagnetic interactions with materials to one loop, including generalized counterterms. We state the renormalization conditions that determine them. Section VI presents our numerical results and compares them to unrenormalized perturbation theory and experiment. Section VII is a summary of the approach. Basic ingredients and some detailed calculations are relegated to four appendixes.

II. THE ELECTROMAGNETIC FREE ENERGY OF A ROUGH AND A FLAT MATERIAL INTERFACE

The present approach is based on Schwinger's low-energy effective field theory [3] for electromagnetism. The partition function in this model is a functional of the local dielectric permittivity tensor $\boldsymbol{\varepsilon}(\zeta_n, \mathbf{x}, z)$ of the material and of an external polarization source $\vec{P}_n(\vec{x}) = \vec{P}(\zeta_n, \vec{x})$. It is the product of contributions from (independent) thermal modes [31] of the electric field to the Matsubara frequency,[1]

$$\zeta_n = 2\pi|n|T \geq 0 \quad \text{for all } n \in \text{Integers}. \tag{1}$$

The partition function formally is given by the functional integral

$$Z_T[\vec{P};\boldsymbol{\varepsilon}] \propto \prod_n \int D[\vec{E}_n] \exp\left\{-\frac{1}{2T}\int d^3x \vec{E}_n^\dagger(\vec{x}) \times \left[\boldsymbol{\varepsilon}(\zeta_n,\vec{x}) + \frac{1}{\zeta_n^2}\nabla\times\nabla\times\right]\vec{E}_n(\vec{x}) + 2T\vec{E}_n(\vec{x})\cdot\vec{P}_n(\vec{x})\right\}. \tag{2}$$

Here $Z_T[\vec{P};\boldsymbol{\varepsilon}]$ is the partition function of QED in axial gauge $A_0 = \Phi = 0$ for a medium with local dielectric permittivity $\boldsymbol{\varepsilon}(\zeta,\vec{x})$. In this gauge, $\vec{E}_n = \zeta\vec{A}_n$, $\vec{B}_n = \nabla\times\vec{A}_n$, and the current source $\vec{j}_n = \zeta\vec{P}_n$.

We consider the standard Casimir configuration of two parallel semi-infinite plates at an average separation a that is much less than their transverse dimension [1]. In the following, we restrict the discussion to the configuration shown in Fig. 1 of two semi-infinite dielectric (metallic) slabs of the same material separated by vacuum, only one of which is rough:

$$\varepsilon_3(\zeta) = 1, \qquad \varepsilon_2(\zeta) = \varepsilon(\zeta) = \varepsilon_1(\zeta). \tag{3}$$

We forego the ability to address lateral Casimir forces, which are finite and vanish if one of the interfaces is flat.

[1]We adopt natural units $\hbar = c = k_B = 1$ and suppress the index n of ζ_n in summations $\sum_n$ over all Matsubara frequencies.

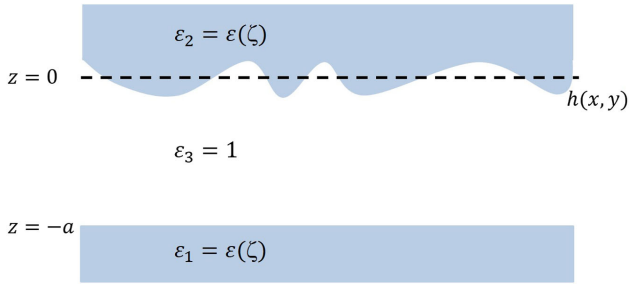

FIG. 1 (color online). Two semi-infinite slabs of the same material separated by vacuum. The low-energy electromagnetic properties of the material are described by a bulk permittivity $\varepsilon(\zeta = i\omega)$ that only depends on the frequency of the electric field. In Cartesian coordinates, the planar interface is at $z = -a$, and the mean separation of the two interfaces is a. The surface of the rough slab is at $z = h(\mathbf{x})$, where $h(\mathbf{x})$ is a profile function that generally depends on both transverse coordinates $\mathbf{x} = (x, y)$. We develop a perturbative expansion valid for $|h(\mathbf{x})| \ll a$ with no restrictions on the profile other than that it be single valued. $h(\mathbf{x})$ in particular need not be as smooth as shown here.

Lateral Casimir forces depend on cross correlations of the two profiles. At separations $a \gg \sigma$, they are small and involve only low momenta. For corrugated plates, they have been computed in Ref. [32]. We here are interested in the effect of profiles on the *normal* Casimir force. The physical interpretation and consistent subtraction of (potentially divergent) contributions will be our main concern.

The rough interface is assumed to be without enclosures, and the deviation from a flat one at $z = 0$ is described by a single-valued function $h(\mathbf{x})$ that satisfies[2]

$$\langle h \rangle := A^{-1} \int_A d\mathbf{r} h(\mathbf{r}) = 0. \tag{4}$$

The point of reference for defining the separation a of the two slabs should be irrelevant. However, a consistent perturbative expansion is feasible only in the absence of so-called tadpole contributions. These vanish if the separation a is such that Eq. (4) holds. Equation (4) in this sense defines the distance a between the interfaces.

When the cross-sectional area A of the slab is taken to be arbitrarily large, boundary effects can be ignored, and the two-point correlation function

$$D_2(\mathbf{x} - \mathbf{y}) = \langle h(\mathbf{x}) h(\mathbf{y}) \rangle := A^{-1} \int_A d\mathbf{r} h(\mathbf{r} + \mathbf{x}) h(\mathbf{r} + \mathbf{y}) \tag{5}$$

is invariant under transverse translations. The roughness variance,

$$\sigma^2 = D_2(0), \tag{6}$$

is a measure for the roughness amplitude.

The dielectric permittivity function $\boldsymbol{\varepsilon}(\varepsilon, \vec{x})$ in this effective low-energy field theory is of the form

$$\begin{aligned} \boldsymbol{\varepsilon}(\zeta, \vec{x}) &= \mathbb{1}[\varepsilon_3(\zeta) + (\varepsilon_2(\zeta) - \varepsilon_3(\zeta))\theta(z - h(\mathbf{x})) \\ &\quad + (\varepsilon_1(\zeta) - \varepsilon_3(\zeta))\theta(-z - a)] \\ &= \mathbf{V}^{\parallel}(\zeta, z) + \mathbf{V}^{h}(\zeta, \mathbf{x}, z), \end{aligned} \tag{7}$$

where

$$\mathbf{V}^{h}(\zeta, \mathbf{x}, z) = \mathbb{1}[(\varepsilon_2(\zeta) - \varepsilon_3(\zeta))(\theta(z - h(\mathbf{x})) - \theta(z))] \tag{8}$$

is the deviation due to the roughness profile $h(\mathbf{x})$ from the dielectric permittivity of a transversely homogeneous medium given by

$$\begin{aligned} \mathbf{V}^{\parallel}(\zeta, z) &= \mathbb{1}[\varepsilon_3(\zeta) + (\varepsilon_2(\zeta) - \varepsilon_3(\zeta))\theta(z) \\ &\quad + (\varepsilon_1(\zeta) - \varepsilon_3(\zeta))\theta(-z - a)] + \delta\mathbf{V}^{h}(\zeta, z). \end{aligned} \tag{9}$$

We shall argue that the counterterm $\delta\mathbf{V}^{h}(\zeta, z)$ to the dielectric permittivity of three flat slabs is necessary for a consistent perturbative expansion in the framework of a low-energy theory. $\delta\mathbf{V}^{h}(\zeta, z)$ depends on gross properties of the profile $h(\mathbf{x})$, but not on the transverse position $\mathbf{x}$, nor on the separation of the two interfaces. This counterterm ensures that the single-interface scattering matrix is reproduced by the low-energy theory. To leading order, $\delta\mathbf{V}^{h}(\zeta, z)$ is proportional to the variance σ^2 of the rough interface. We are thus calculating the perturbative expansion for the rough interface at $z = 0$ about an effective $\mathbf{x}$-independent (bare) permittivity,

$$\boldsymbol{\varepsilon}_{\text{eff}}(\zeta, z) = \mathbb{1}\varepsilon_2(\zeta)\theta(z) + \mathbb{1}\varepsilon_3(\zeta)\theta(-z) + \delta\mathbf{V}^{h}(\zeta, z). \tag{10}$$

$\delta\mathbf{V}^{h}(\zeta, z)$ has support near the surface at $z \sim 0$ only.[3] To approximate scattering off a rough interface by an effective $\boldsymbol{\varepsilon}_{\text{eff}}(\zeta, z)$ is a conceptually appealing idea and not new [33,34]. We develop a consistent low-energy approach in which this is realized perturbatively. Contrary to commonly used *Ansätze* for the effective $\boldsymbol{\varepsilon}_{\text{eff}}$, $\delta\mathbf{V}^{h}(\zeta, z)$ generally is not isotropic.

The inherent limitations of the effective low-energy description derive from the fact that electromagnetic interaction with matter is encoded in the permittivity function. They are not restricted to a perturbative analysis. The dimensionless permittivity $\varepsilon(\zeta) = \varepsilon(\zeta/\omega_p)$ depends implicitly on a scale that can be identified with the plasma frequency ω_p of the material. For momentum or energy

[2]A Cartesian coordinate system with the z axis normal to the plates is used to describe this system. We use bold type $\mathbf{x} = (x, y)$ for two-dimensional vectors perpendicular to the z axis, whereas $\vec{v}$ denotes an ordinary three-dimensional vector.

[3]To first order in the variance, we find in Eq. (43) of Sec. III that $\delta\mathbf{V}^{h}(\zeta, z) \propto \delta(z)$—that is, $\boldsymbol{\varepsilon}_{\text{eff}}(\zeta, z)$ differs from that of an interface by the insertion of an arbitrarily thin plate.

transfers (or temperatures) that are much larger than ω_p, the effective low-energy theory of Eq. (2) fails to incorporate nonlinear effects or to account for the creation of free charges. The *Ansatz* that the permittivity does not depend on the profile, furthermore, is incorrect at wavelengths comparable to the plasma wavelength $l_p = 2\pi/\omega_p$, and a description in terms of the bulk permittivity of the homogeneous material is not warranted within the plasma-skin depth of order l_p. For gold surfaces commonly used, $\omega_p \sim 0.046\ \text{nm}^{-1} \sim 9$ eV. The low-energy description of electromagnetic interactions with such materials by Eq. (7), therefore, is already questionable at wave numbers $q \gg \omega_p \sim 0.046\ \text{nm}^{-1}$ that resolve less than 20 nm, or about 200 gold atoms. We will find that roughness corrections to the Casimir force with correlation lengths $l_c \lesssim 1/\omega_p$ depend on momentum transfers $q \gg \omega_p$ that are inadequately described by the low-energy theory. The conservative approach is to use the effective low-energy theory to only compute roughness corrections with $l_c \gg 1/\omega_p \sim 20$ nm, a regime where the PFA generally is quite accurate. We improve on this by introducing phenomenological input.

It is interesting in this regard that many comparisons of theory with experiments in the literature are for correlation lengths $l_c \sim 25\ \text{nm} \sim 1/\omega_{p\,\text{Au}}$. The unimproved theory is highly sensitive to l_c in this regime and (re)produces large variations with only small changes in parameters. Roughness corrections computed with this unimproved model for such short correlation lengths are uncontrolled, and in fact physically untenable [22].

We here compute roughness corrections that are consistent with the low-energy effective model by using low-energy (experimental) data to systematically subtract and correct high-momentum contributions to the loop expansion. The method is quite general [26,27] and has been successfully applied to low-energy effective field theories as diverse as chiral perturbation theory [28] and (quantum) gravity [27]. In our case, it yields a consistent expansion in σ/a for any value of $0 < l_c\omega_p < \infty$ at the expense that the reflection of electromagnetic radiation perpendicular to the rough plate has to either be measured or be reliably modeled.

A. The Green's function and Casimir energy of two parallel flat interfaces

Schwinger obtained the free energy and the response to an external polarization source $\vec{P}_n(\mathbf{x}, z) = \vec{P}(\mathbf{x}, z; \zeta_n)$ for three parallel slabs in the framework of the low-energy effective field theory given by Eq. (2). The free energy in this case is [3]

$$\mathcal{F}_T^{\parallel}(a, \vec{P}) = F_T^{\parallel}(a) - \frac{T^2}{2}\sum_n \{\vec{P}_n|\mathbf{G}^{\parallel(n)}|\vec{P}_n\}. \quad (11)$$

Here $F_T^{\parallel}$ is the well-known Casimir free energy for three parallel slabs,

$$F_T^{\parallel}(a) = \frac{AT}{2}\sum_n \int \frac{d\mathbf{k}}{(2\pi)^2}[\ln(1 - r_1 r_2 e^{-2\kappa_3 a}) + \ln(1 - \bar{r}_1 \bar{r}_2 e^{-2\kappa_3 a})], \quad (12)$$

where the reflection coefficients at the ith interface of area A for the TE and TM modes are

$$r_i = r_i^{\text{TE}} = \frac{\kappa_3 - \kappa_i}{\kappa_3 + \kappa_i} \quad \text{and} \quad \bar{r}_i = r_i^{\text{TM}} = \frac{\bar{\kappa}_3 - \bar{\kappa}_i}{\bar{\kappa}_3 + \bar{\kappa}_i}, \quad \text{with}$$
$$\kappa_i = \sqrt{\mathbf{k}^2 + \zeta^2 \varepsilon_i(\zeta)} \quad \text{and} \quad \bar{\kappa}_i = \frac{\kappa_i}{\varepsilon_i(\zeta)}. \quad (13)$$

The response to the nth Matsubara mode of an external source of polarization is

$$\{\vec{P}_n|\mathbf{G}^{\parallel(n)}|\vec{P}_n\} = \int dz dz' d\mathbf{x} d\mathbf{y} \vec{P}_n^{\dagger}(\mathbf{x}, z) \cdot \mathbf{G}^{\parallel}(\mathbf{x}, z, \mathbf{y}, z'; \zeta_n, a) \cdot \vec{P}_n(\mathbf{y}, z) \quad (14a)$$

$$= \int \frac{d\mathbf{k}}{(2\pi)^2} dz dz' \vec{P}_n^{\dagger}(\mathbf{k}, z) \cdot \mathbf{G}^{\parallel}(\mathbf{k}, z, z'; \zeta_n, a) \cdot \vec{P}_n(\mathbf{k}, z). \quad (14b)$$

$\mathbf{G}^{\parallel}$ in Eq. (14a) is the Green's dyadic solving,[4]

$$\left[\mathbf{V}^{\parallel}(\zeta, z) + \frac{1}{\zeta^2}\nabla\times\nabla\times\right]\mathbf{G}^{\parallel}(\mathbf{x}, z, \mathbf{y}, z'; \zeta, a) = \mathbb{1}\delta(z - z')\delta(\mathbf{x} - \mathbf{y}). \quad (15)$$

Due to translational invariance in transverse directions, $\mathbf{G}^{\parallel}(\mathbf{x}, z, \mathbf{y}, z'; \zeta, a)$ is a function of $\mathbf{x} - \mathbf{y}$, and the Fourier representations in Eq. (14b) are

$$\mathbf{G}^{\parallel}(\mathbf{x}, z, \mathbf{y}, z'; \zeta, a) = \int \frac{d\mathbf{k}}{(2\pi)^2} e^{i\mathbf{k}(\mathbf{x}-\mathbf{y})} \mathbf{G}^{\parallel}(\mathbf{k}, z, z'; \zeta, a)$$
$$\text{and} \quad \vec{P}_n(\mathbf{k}, z) = \int d\mathbf{x} e^{-i\mathbf{k}\mathbf{x}} \vec{P}_n(\mathbf{x}, z). \quad (16)$$

$\mathbf{G}^{\parallel}$ can be decomposed into a single-interface Green's dyadic $\mathbf{G}^{|}(\mathbf{x} - \mathbf{y}, z, z'; \zeta) = \mathbf{G}^{\parallel}(\mathbf{x}, z, \mathbf{y}, z'; \zeta, a \to \infty)$, where the second interface has been removed, and the correction $\mathbf{G}^{|a|}(\mathbf{x} - \mathbf{y}, z, z'; \zeta, a)$ is due to the presence of a second flat interface at mean separation a. In momentum space, the latter vanishes exponentially for $a \to \infty$:

$$\mathbf{G}^{\parallel}(\mathbf{k}, z, z'; \zeta, a) = \mathbf{G}^{|}(\mathbf{k}, z, z'; \zeta) + \mathbf{G}^{|a|}(\mathbf{k}, z, z'; \zeta). \quad (17)$$

Explicit expressions for components of $\mathbf{G}^{|}(\mathbf{k}, z, z'; \zeta)$ and $\mathbf{G}^{|a|}(\mathbf{k}, z, z'; \zeta)$ when z and z' are in slab #2 or slab #3 are collected in Appendix A.

[4] $\mathbf{G}^{\parallel}(\zeta)$ is related to Schwinger's [3] dyadic $\mathbf{\Gamma}(\omega)$ at angular frequency ω by $\mathbf{G}^{\parallel}(\zeta) = -\mathbf{\Gamma}(i\omega)$.

B. Perturbative roughness correction to the Casimir free energy: Green's function formalism

A straightforward perturbative expansion in the roughness potential $\mathbf{V}^h$ is possible only for media with $\varepsilon_2 - \varepsilon_3 \ll 1$. Since the Casimir free energy itself is rather small, roughness corrections are not very important in this weak coupling scenario. However, the support of $\mathbf{V}^h$ is restricted to $|z| \leq \max_{\mathbf{x}} |h(\mathbf{x})| \sim \sigma \ll a$, and a perturbative expansion in σ/a may exist even for media whose permittivity is rather large. This expansion, in fact, is possible even for ideal metals.

The part of the free energy that captures the dependence on the average separation a of two interfaces is by definition the Casimir free energy due to their interaction.[5] In terms of the Green's dyadic $\mathbf{G}^{\parallel}$ of three parallel slabs satisfying Eq. (15), the full Green's dyadic $\mathbf{G}$ for the combination of a rough and a flat interface formally is the solution of

$$[\mathbb{1} + \mathbf{V}\mathbf{G}^{\parallel}]\mathbf{G}^{\parallel -1}\mathbf{G} = \mathbb{1}, \tag{18}$$

with

$$\begin{aligned}\mathbf{V} &= \mathbf{V}^h + \delta\mathbf{V}^h \\ &= \mathbb{1}(\varepsilon(\zeta) - 1)(\theta(z - h(\mathbf{x})) - \theta(z)) + \delta\mathbf{V}^h(\zeta, z).\end{aligned} \tag{19}$$

The change in free energy due to the roughness of one interface, therefore, is [32,35]

$$\Delta F_T[h,a] = -\frac{1}{2}\mathrm{Tr}\ln(\mathbf{G}^{\parallel -1}\mathbf{G}) = \frac{1}{2}\mathrm{Tr}\ln(\mathbb{1} + \mathbf{V}\mathbf{G}^{\parallel}), \tag{20}$$

where the trace includes a summation over Matsubara frequencies and over a complete set of scattering states. The expression in Eq. (20) is rather formal, because it includes the change in free energy due to roughness in the absence of the second (flat) interface. This infinite single-body contribution to the free energy does not depend on the mean separation a. Subtracting from $\Delta F_T[h,a]$ its value when the two interfaces are infinitely far apart gives the correction to the Casimir free energy due to the roughness of an interface as

$$\begin{aligned}\Delta F_T^{\mathrm{Cas}}[h,a] &:= \Delta F_T[h,a] - \Delta F_T[h,\infty] \\ &= \frac{1}{2}\mathrm{Tr}\ln(\mathbb{1} + \mathbf{V}\mathbf{G}^{\parallel}) - \frac{1}{2}\mathrm{Tr}\ln(\mathbb{1} + \mathbf{V}\mathbf{G}^{|}) \\ &= \frac{1}{2}\mathrm{Tr}\ln(\mathbb{1} + \mathbf{T}^h\mathbf{G}^{|a|}),\end{aligned} \tag{21}$$

where

$$\mathbf{T}^h = \mathbf{V} - \mathbf{V}\mathbf{G}^{|}\mathbf{T}^h \tag{22}$$

is the formal scattering matrix due to the roughness potential $\mathbf{V}$. $\mathbf{T}^h$ does not depend on the separation a and describes scattering due to roughness in the *absence* of the second (flat) interface. Since high momenta are exponentially suppressed in $\mathbf{G}^{|a|}$, the Volterra series of $\Delta F_T^{\mathrm{Cas}}[h,a]$ in powers of $\mathbf{T}^h$,

$$\begin{aligned}\Delta F_T^{\mathrm{Cas}}[h,a] &= \frac{1}{2}\mathrm{Tr}\ln(\mathbb{1} + \mathbf{T}^h\mathbf{G}^{|a|}) \\ &\sim \frac{1}{2}\mathrm{Tr}[\mathbf{T}^h\mathbf{G}^{|a|} - \frac{1}{2}\mathbf{T}^h\mathbf{G}^{|a|}\mathbf{T}^h\mathbf{G}^{|a|} + \cdots,\end{aligned} \tag{23}$$

converges when the norm of $\mathbf{T}^h\mathbf{G}^{|a|}$ is bounded and sufficiently small.

III. THE ROUGHNESS SCATTERING MATRIX $\mathbf{T}^h$

Noting that the component $G^{|}_{zz}(\mathbf{k}, z, \mathbf{y}, z'; \zeta)$ in Eq. (A1) includes a δ-function singularity, Eq. (22) can be rewritten

$$\mathbf{T}^h = \tilde{\mathbf{V}} - \tilde{\mathbf{V}}\tilde{\mathbf{G}}^{|}\mathbf{T}^h \tag{24}$$

in terms of the Green's dyadic $\tilde{\mathbf{G}}^{|}$ with Fourier components,

$$\tilde{\mathbf{G}}^{|}(\mathbf{k}, z, z'; \zeta) = \mathbf{G}^{|}(\mathbf{k}, z, z'; \zeta) - \mathrm{diag}\left(0, 0, \frac{1}{\varepsilon_z}\delta(z - z')\right), \tag{25}$$

and a new potential $\tilde{\mathbf{V}}$. $\tilde{\mathbf{G}}$ is devoid of δ-function singularities (but not continuous at $z = 0$) with the components given in Eq. (A7). To order σ^2, the potential $\tilde{\mathbf{V}}$ is

$$\begin{aligned}\tilde{\mathbf{V}}(\mathbf{x}, z; \zeta) &= \tilde{\mathbf{V}}^h(\zeta, \mathbf{x}, z) + \delta\tilde{\mathbf{V}}^h(\zeta, z) \\ &= (\varepsilon - 1)[\theta(z - h(\mathbf{x})) - \theta(z)]\mathrm{diag}[1, 1, \varepsilon\theta(z) \\ &\quad + \theta(-z)/\varepsilon] + \delta\tilde{\mathbf{V}}^h(\zeta, z).\end{aligned} \tag{26}$$

The reformulation of Eq. (22) in the form of Eq. (24) resums local contributions of the same order in h. It allows the formulations of a consistent perturbative expansion in σ even in the ideal metal limit $\varepsilon(\zeta) \to \infty$. Just as for $\mathbf{V}$, the support of $\tilde{\mathbf{V}}$ is restricted to the interval $|z| < \max_{\mathbf{x}}|h(\mathbf{x})| \sim \sigma$ only. Since $\tilde{\mathbf{G}}^{|}$ is free of ultralocal δ-function singularities, contributions to $\mathbf{T}^h$ of nth order in $\tilde{\mathbf{V}}$ are at least of nth order in the standard deviation σ of the profile h.

To second order in σ, we need only consider the first two terms of the Volterra series,

$$\mathbf{T}^h \approx \tilde{\mathbf{V}} - \tilde{\mathbf{V}}\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}} \approx \tilde{\mathbf{V}} - \tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}}^h = \mathbf{T}^{(1)} + \mathbf{T}^{(2)}, \tag{27}$$

since the counterterm potential $\delta\tilde{\mathbf{V}}^h$ is itself of order σ^2 (as will be seen). The second-order contribution $\mathbf{T}^{(2)}$ of Eq. (27) is at least of order σ^2, and its integrated expectation to this order is

[5] It vanishes in the limit $a \to \infty$. At $T = 0$, this "free energy" is the Casimir energy and need not vanish.

$$\mathbf{t}^{(2)}(\mathbf{x}-\mathbf{y},\zeta) \coloneqq \left\langle \int dz dz' \mathbf{T}^{(2)}(\mathbf{x},z,\mathbf{y},z';\zeta) \right\rangle = -\left\langle \int dz dz' \tilde{\mathbf{V}}^h(\mathbf{x},z;\zeta)\tilde{\mathbf{G}}^|(\mathbf{x}-\mathbf{y},z,z';\zeta)\tilde{\mathbf{V}}^h(\mathbf{y},z';\zeta) \right\rangle + \mathcal{O}(\sigma^3). \quad (28)$$

Because $\int dz \tilde{\mathbf{V}}^h(\mathbf{x},z,\zeta)$ already is of order σ, the Fourier components of $\mathbf{t}^{(2)}$ are[6]

$$\begin{aligned}
t^{(2)}_{xx}(k,\zeta) &= -(\varepsilon-1)^2 \int \frac{d\mathbf{k}'}{(2\pi)^2}\left(\frac{\kappa'\kappa'_\varepsilon \cos^2\theta}{\kappa'_\varepsilon+\varepsilon\kappa'} + \frac{\zeta^2\sin^2\theta}{\kappa'_\varepsilon+\kappa'}\right) D(q) + \mathcal{O}(\sigma^3),\\
t^{(2)}_{yy}(k,\zeta) &= -(\varepsilon-1)^2 \int \frac{d\mathbf{k}'}{(2\pi)^2}\left(\frac{\kappa'\kappa'_\varepsilon \sin^2\theta}{\kappa'_\varepsilon+\varepsilon\kappa'} + \frac{\zeta^2\cos^2\theta}{\kappa'_\varepsilon+\kappa'}\right) D(q) + \mathcal{O}(\sigma^3),\\
t^{(2)}_{zz}(k,\zeta) &= -(\varepsilon-1)^2 \int \frac{d\mathbf{k}'}{(2\pi)^2}\frac{-k'^2}{\varepsilon\kappa'+\kappa'_\varepsilon}\left(\varepsilon D_{++}(q) + D_{-+}(q) + D_{+-}(q) + \frac{1}{\varepsilon}D_{--}(q)\right) + \mathcal{O}(\sigma^3),\\
t^{(2)}_{xz}(k,\zeta) &= -(\varepsilon-1)^2 \int \frac{d\mathbf{k}'}{(2\pi)^2}\frac{ik'\cos\theta}{\kappa'_\varepsilon+\varepsilon\kappa'}(\kappa'_\varepsilon D_{++}(q) + \bar{\kappa}'_\varepsilon D_{+-}(q) - \varepsilon\kappa' D_{-+}(q) - \kappa' D_{--}(q)) + \mathcal{O}(\sigma^3),\\
t^{(2)}_{zx}(k,\zeta) &= -(\varepsilon-1)^2 \int \frac{d\mathbf{k}'}{(2\pi)^2}\frac{-ik'\cos\theta}{\kappa'_\varepsilon+\varepsilon\kappa'}(\kappa'_\varepsilon D_{++}(q) - \varepsilon\kappa' D_{+-}(q) + \bar{\kappa}'_\varepsilon D_{-+}(q) - \kappa' D_{--}(q)) + \mathcal{O}(\sigma^3),
\end{aligned} \quad (29)$$

with $q^2 = (\mathbf{k}-\mathbf{k}')^2 = k^2 + k'^2 - 2kk'\cos\theta$. Since the $\tilde{G}^|_{zx}$ and $\tilde{G}^|_{zz}$ components of the dyadic (see Appendix A) are discontinuous at $z=0$, one has to separately consider correlators of positive and negative components of the roughness profile in Eq. (29). With $h_\pm(\mathbf{x}) = h(\mathbf{x})\theta(\pm h(\mathbf{x}))$, these signed correlators are

$$\begin{aligned}
D_{\pm+}(q) = D_{\mp-}(q) &\coloneqq \int d\mathbf{x} e^{i\mathbf{q}(\mathbf{x}-\mathbf{y})}\langle h_\pm(\mathbf{x})h_+(\mathbf{y})\rangle,\\
D(q) = 2D_{++}(q) + 2D_{+-}(q) &= \int d\mathbf{x} e^{i\mathbf{q}(\mathbf{x}-\mathbf{y})}\langle h(\mathbf{x})h(\mathbf{y})\rangle.
\end{aligned} \quad (30)$$

$D(q)$ is the Fourier transform of the two-point correlation function $D_2(\mathbf{x}-\mathbf{y})$ of Eq. (60). As shown in Appendix B, the signed correlators for a Gaussian generating functional of roughness correlations are also related to the two-point correlator D_2 as

$$\begin{aligned}
\langle h_+(\mathbf{x})h_+(\mathbf{y})\rangle &= \langle h_-(\mathbf{x})h_-(\mathbf{y})\rangle = \frac{\sigma^2}{2\pi}(\sin\phi + (\pi-\phi)\cos\phi),\\
\langle h_+(\mathbf{x})h_-(\mathbf{y})\rangle &= \langle h_-(\mathbf{x})h_+(\mathbf{y})\rangle = \frac{\sigma^2}{2\pi}(\phi\cos\phi - \sin\phi),\\
&\text{with } \ 0 \le \cos\phi = D_2(\mathbf{x}-\mathbf{y})/D_2(0) \le 1
\end{aligned} \quad (31)$$

for a roughness correlation function $D_2(r)$ that is positive and monotonically decreasing with $r = |\mathbf{x}-\mathbf{y}|$. The signed correlators do not vanish and approach $\pm\sigma^2/(2\pi)$ for $r\to\infty$ if $D_2(r\sim\infty)\sim 0$. At small separations, $r = |\mathbf{x}-\mathbf{y}| \ll l_c$, $\cos\phi = D_2(r)/D_2(0) \sim 1 - \beta r^\alpha$. Thus, $\phi \propto r^{\alpha/2}$ for $r\sim 0$ with an exponent $\alpha > 0$. The expressions of Eq. (31) for small ϕ then imply the behavior

$$\begin{aligned}
&\langle h_+(\mathbf{x})h_+(\mathbf{y})\rangle = \langle h_-(\mathbf{x})h_-(\mathbf{y})\rangle \sim \frac{1}{2}D_2(r);\\
&\langle h_+(\mathbf{x})h_-(\mathbf{y})\rangle = \langle h_-(\mathbf{x})h_+(\mathbf{y})\rangle \sim -\frac{\sigma^2}{6\pi}(2\beta r^\alpha)^{3/2} \ \text{for } r \ll l_c.
\end{aligned} \quad (32)$$

After Fourier transformation, the asymptotic behavior at large momenta $q l_c \gg 1$ of $D_{++}(q) = D_{--}(q)$ is the same as that of $\frac{1}{2}D(q \gg 1/l_c)$, whereas the mixed correlations $D_{+-}(q) = D_{-+}(q)$ fall off more rapidly.

For $l_c \ll 1/\omega_p$, high-momentum contributions are appreciable or even dominate the one-loop corrections to the diagonal components of the scattering matrix in Eq. (29). For example,

$$\begin{aligned}
t^{(2)}_{xx}(0,\zeta) &= -(\varepsilon-1)^2\int_0^\infty \frac{kdk}{4\pi}\left(\frac{\kappa\kappa_\varepsilon}{\kappa_\varepsilon+\varepsilon\kappa} + \frac{\zeta^2}{\kappa_\varepsilon+\kappa}\right)D(k)\\
&\xrightarrow{l_c\sim 0} -\frac{(\varepsilon-1)^2}{1+\varepsilon}\int_0^\infty \frac{kdk}{4\pi}kD(k).
\end{aligned} \quad (33)$$

Whether or not loop integrals like Eq. (33) diverge depends on the roughness correlation function. For Gaussian correlations, the integral converges,

$$\begin{aligned}
&\langle h(\mathbf{x})h(\mathbf{y})\rangle = \sigma^2 e^{-\frac{1}{2}(\mathbf{x}-\mathbf{y})^2/l_c^2} \Rightarrow D_{\text{Gauss}}(q) = 2\pi\sigma^2 l_c^2 e^{-\frac{1}{2}q^2 l_c^2},\\
&\text{with} \quad \int_0^\infty \frac{kdk}{4\pi}kD_{\text{Gauss}}(k) = \frac{\sigma^2}{2l_c}\sqrt{\frac{\pi}{2}},
\end{aligned} \quad (34)$$

but the roughness "correction" becomes (arbitrarily) large for $l_c \sim 0$. This invalidates the perturbative expansion in σ/a and, for sufficiently small l_c, violates unitarity. It furthermore is unphysical that roughness corrections to the scattering matrix with profiles of fixed variance become arbitrarily large as $l_c \to 0$.

For a scalar field and Gaussian roughness correlation, higher orders in the loop expansion are of the same order in σ/l_c in this limit [22]. Assuming the scalar model is valid at all energy scales, we resummed the leading σ/l_c contributions to the

[6] $\mathbf{k} = (k,0,0)$ here defines the positive x axis, and (k',θ) are polar coordinates of $\mathbf{k}'$. Note that a (randomly) rough profile preserves translational (and rotational) invariance on average. The *average* scattering matrix of Eq. (28), therefore, is diagonal in transverse momentum space.

scalar Casimir energy and found that they amount to a change in the effective separation $\Delta a \sim \sigma^2/l_c$ of the two interfaces.

However, the effective low-energy electromagnetic theory of Eq. (2) evidently is not valid for momenta that far exceed the plasma frequency ω_p. One furthermore is not assured that summing incorrect higher-loop contributions in this effective low-energy theory improves the situation. We therefore will not follow that line and will proceed differently in this case.

For correlation functions with nonvanishing slope at $r = |\mathbf{x}-\mathbf{y}| = 0$—that is, $D_2'(r=0) \neq 0$—the situation is even more serious. For instance, the two-dimensional Fourier transform of an exponential correlation function,

$$\langle h(\mathbf{x})h(\mathbf{y})\rangle = \sigma^2 e^{-|(\mathbf{x}-\mathbf{y})|/l_c} \Rightarrow D_{\mathrm{Exp}}(q) = \frac{2\pi\sigma^2 l_c^2}{(1+q^2 l_c^2)^{3/2}}, \tag{35}$$

decays as a power law proportional to q^{-3} at large momenta. The integral in Eq. (33) and other (diagonal) components of the roughness correction $\mathbf{t}^{(2)}$ in Eq. (29) to the single-interface scattering matrix in this case are logarithmic UV divergent for *any* correlation length $l_c > 0$.

Experiment [19] does not distinguish Gaussian roughness correlations,[7] and roughness profiles with correlation lengths $l_c\omega_p \ll 1$ are readily manufactured. Restricting the model to a particular form for the roughness correlation would not address the fact that the effective low-energy theory does not describe high-momentum contributions to loop integrals correctly.

From a practical point of view, the problem is that roughness corrections to the Casimir free energy and other low-energy observables are exceptionally *sensitive* to high-frequency components of the profile because $\mathbf{G}^{|}(k \sim \infty, 0, 0; \zeta) \sim k$ at large momenta. Figure 2 depicts typical roughness profiles to three different correlation functions with the same correlation length and variance: (a) exponential as in Eq. (35), (b) Gaussian as in Eq. (34), and (c) rational as $D_{\mathrm{Rational}}(r) = \sigma^2/(1+(r/l_c)^2)^2$. It is evident from Fig. 2 that the three profiles differ only in their high-frequency components. However, to leading order in the variance, corrections to the low-energy scattering matrix

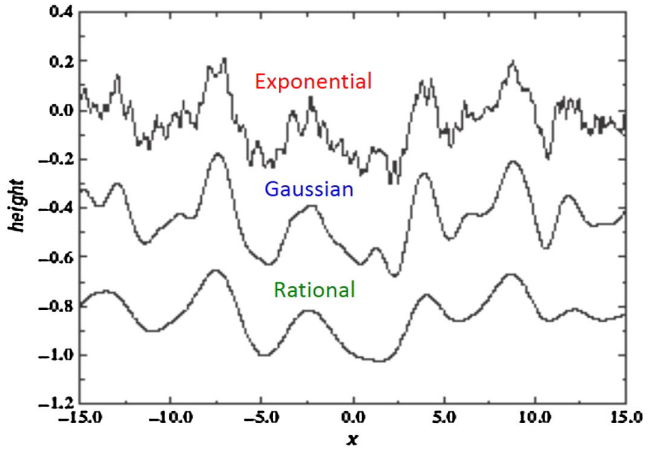

FIG. 2 (color online). Typical cross sections of two-dimensional profiles with different correlations (reproduced from Ref. [36]). From the top: Profile with the exponential correlation $D_{\mathrm{Exp}}(q)$ of Eq. (35), profile with the Gaussian correlation $D_{\mathrm{Gauss}}(q)$ of Eq. (34), profile with a rational correlation (see text). The correlation length and variance are the same for all three profiles. For clarity, the average height of the profiles differs by −0.4. Units are arbitrary. Note that only high-frequency components of the profiles differ significantly.

are extremely different for the three types of profiles. The roughness correction diverges in the exponential case (a) but is finite for profiles (b) and (c). This sensitivity can be traced to the UV behavior of the one-loop integrands like that of Eq. (33). It is unphysical and an artifact of taking the low-energy effective theory beyond its limits.

Analogous difficulties arise in any nonrenormalizable low-energy effective field theory [27,28], and we here resort to a similar cure: whereas high momenta may dominate loop corrections to the scattering matrix, they generally are sufficiently suppressed in differences thereof. Differences of elements of the scattering matrix often can be reliably estimated within the framework of the low-energy effective field theory. However, phenomenological input is required to determine high-momentum contributions to loop integrals that are beyond the reach of the low-energy theory.

One, for instance, can rewrite $t_{xx}^{(2)}(\mathbf{k},\zeta)$ of Eq. (29) in the form

$$\begin{aligned} t_{xx}^{(2)}(\mathbf{k},\zeta) &= t_{xx}^{(2)}(0,\zeta) + (t_{xx}^{(2)}(\mathbf{k},\zeta) - t_{xx}^{(2)}(0,\zeta)) \\ &= t_{xx}^{(2)}(0,\zeta) - (\varepsilon-1)^2 \int \frac{d\mathbf{k}'}{(2\pi)^2} \left(\frac{\kappa'\kappa_\varepsilon' \cos^2\theta}{\kappa_\varepsilon' + \varepsilon\kappa'} + \frac{\zeta^2 \sin^2\theta}{\kappa_\varepsilon' + \kappa'} \right) (D(|\mathbf{k}'-\mathbf{k}|) - D(k')) \\ &= t_{xx}^{(2)}(0,\zeta) - (\varepsilon-1)^2 \int \frac{d\mathbf{q}}{(2\pi)^2} \left(\frac{\kappa'\kappa_\varepsilon'(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')^2}{\kappa_\varepsilon' + \varepsilon\kappa'} + \frac{\zeta^2(1-(\hat{\mathbf{k}}\cdot\hat{\mathbf{k}}')^2)}{\kappa_\varepsilon' + \kappa'} - \frac{\kappa\kappa_\varepsilon(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}})^2}{\kappa_\varepsilon + \varepsilon\kappa} - \frac{\zeta^2(1-(\hat{\mathbf{k}}\cdot\hat{\mathbf{q}})^2)}{\kappa_\varepsilon + \kappa} \right) D(q), \end{aligned} \tag{36}$$

[7] $D_{\mathrm{Gauss}}(q) = D_\infty(q)$ and $D_{\mathrm{Exp}}(q) = D_{1/2}(q)$ in the class of L_1 correlations $\{D_s(q) := 2\pi\sigma^2 l_c^2 (1+\frac{q^2 l_c^2}{2s})^{-s-1}$, with $s>0\}$. The corresponding coordinate-space correlation functions are $D_s(r) = \sigma^2 \frac{2(r\sqrt{2s}/l_c)^s}{2^s\Gamma[s]} K_s(r\sqrt{2s}/l_c)$. Reference [12] uses a correlation in this affine class with $s = 0.9$ for which the loop integral converges but is sensitive to contributions from high momenta.

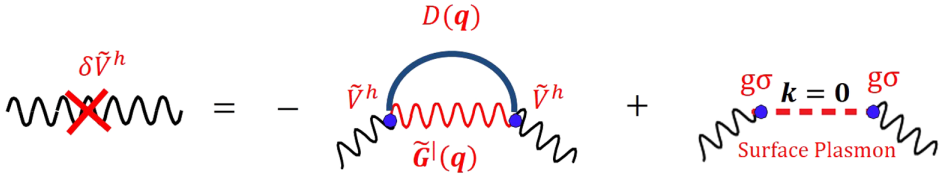

FIG. 3 (color online). The counterpotential $\tilde{\mathbf{V}}^h$ includes two contributions of order σ^2. It subtracts the one-loop contribution to the average scattering matrix at vanishing (transverse) momentum and replaces it with the phenomenological one. The latter is modeled by the tree-level plasmon contribution at vanishing transverse momentum. The plasmon couples to radiation due to the roughness of the surface only, and its coupling $g^2\sigma^2$ to this order is proportional to the variance of the roughness profile. The plasmon propagator (dashed) is the one-interface Green's function $\tilde{\mathbf{G}}(z = z' = k = 0)$. We show in the text that $g^2(\zeta/\omega_p, l_c\omega_p) = 1$ at low frequencies.

where $\mathbf{q} = \mathbf{k}' - \mathbf{k}$, $\kappa'_\varepsilon = \sqrt{(\mathbf{k}+\mathbf{q})^2 + \zeta^2\varepsilon(\zeta)}$ and $\kappa_\varepsilon = \sqrt{\mathbf{q}^2 + \zeta^2\varepsilon(\zeta)}$ in the last expression. The one-loop correction to $t^{(2)}_{xx}(0,\zeta)$ in Eq. (36) converges for any $D(q)$ for which

$$\langle h^2(\mathbf{x})\rangle = \int_0^\infty \frac{qdq}{2\pi} D(q) = \sigma^2 < \infty. \tag{37}$$

More importantly, the correction to $t^{(2)}_{xx}(0,\zeta)$ in Eq. (36) is of order $(k\sigma)^2$ and is thus small at low transverse momenta for *any* correlation length l_c of the profile. This correction to $t^{(2)}_{xx}(0,\zeta)$ thus is reliably computed in the framework of the low-energy theory.

It remains to estimate $\mathbf{t}^{(2)}(0,\zeta)$. This is the correction due to roughness to the (analytically continued) scattering matrix of an electromagnetic wave of frequency $\omega = i\zeta$ incident perpendicular to the rough plate. $\mathbf{t}^{(2)}(0,\zeta)$ is a single-interface low-energy characteristic that, at least in principle, can be derived from ellipsometric measurements of the rough interface. Instead of directly incorporating such experimental data, we here model the corrections of order σ^2 to the low-energy scattering matrix by the coupling to surface plasmons induced by roughness. We determine the coupling by demanding that this phenomenological description of $\mathbf{t}^{(2)}(0,\zeta)$ be consistent with the low-energy field theory in the limit of large correlation length and that the ideal metal limit exist at any correlation length.

Roughness couples electromagnetic radiation to surface plasmons [37]. At low transverse wave numbers, this coupling is of the order of the rms roughness σ. To order σ^2, the corresponding tree-level correction to the scattering matrix is schematically shown in Fig. 3. The diagram depicts the creation, propagation, and subsequent annihilation of a surface plasmon by an incident electromagnetic wave.

For $\mathbf{k} \to 0$, a surface plasmon on the interface of a flat plate at $z = 0$ propagates with the dyadic

$$\mathbf{G}_{\text{plasmon}}(\mathbf{k}\sim 0;\zeta) = \tilde{\mathbf{G}}^{|}(k = 0, z = z' = 0;\zeta) \sim \frac{\zeta}{1+\sqrt{\varepsilon(\zeta)}}\,\text{diag}(1,1,0). \tag{38}$$

To second order in σ, the correction $\mathbf{t}^{(2)}(0,\zeta)$ to the scattering matrix at vanishing momentum transfer from surface plasmons thus is

$$\mathbf{t}^{(2)}(\mathbf{k}=0,\zeta) \approx \mathbf{t}^{(2)}_{\text{Plasmon}}(\mathbf{k}=0,\zeta) = -\sigma^2 g^2 \frac{\zeta(\varepsilon(\zeta)-1)^2}{1+\sqrt{\varepsilon(\zeta)}}\,\text{diag}(1,1,0), \tag{39}$$

where $g(\zeta/\omega_p; l_c\omega_p)$ is a dimensionless coupling that depends only on the frequency of the plane wave incident perpendicular to the rough plate. The coupling $g(\zeta/\omega_p; l_c\omega_p)$ in general is not calculable within this low-energy effective model and has to be determined phenomenologically. We argue below that $g^2 \sim 1$ at low energies.

Since $g(\zeta/\omega_p; l_c\omega_p)$ is a phenomenological *function* rather than just a constant, one could have directly modeled $\mathbf{t}^{(2)}(\mathbf{k}=0,\zeta)$. However, the *Ansatz* of Eq. (39) is consistent with the low-energy scattering theory in the sense that roughness correlation functions for large correlation lengths $l_c\omega_p \gg 1$ approach representations of the δ distribution[8]

$$\lim_{l_c\to\infty} D(\mathbf{q}; l_c) = (2\pi)^2\sigma^2\delta(\mathbf{q}). \tag{40}$$

Loop integrals in the limit $l_c \to \infty$ become trivial, and furthermore, involve only momenta $q \ll \omega_p$. Predictions of the low-energy theory, therefore, should be reliable in the limit $l_c \to \infty$. Evaluating the loop integrals of Eq. (29) for $k \to 0$ using Eq. (40) and comparing with the plasmon contribution of Eq. (39), this requires that

$$g(\zeta/\omega_p, l_c\omega_p \sim \infty) = 1. \tag{41}$$

We will find that Eq. (41) ensures not only consistency, but also the existence of an ideal metal limit. It in addition ensures that the PFA to the Casimir free energy is recovered in the limit $l_c\omega_p \to \infty$.

At finite $l_c\omega_p \lesssim 1$, the coupling $g(\zeta/\omega_p, l_c\omega_p \lesssim 1)$ in principle has to be determined phenomenologically. However, the coupling is severely constrained if we impose

[8]On the space of measurable L^0 test functions. The *subtracted* loop integrand is in this class.

some theoretical requirements. Since the range of frequencies ζ that contribute to the Casimir energy satisfy $\zeta a \lesssim 1 \ll \omega_p a$, and the plasmon coupling does not diverge at low frequencies, we in the following ignore the ζ dependence of $g(\zeta/\omega_p, l_c\omega_p)$, and for low frequencies approximate

$$g(\zeta/\omega_p, l_c\omega_p) \sim g(l_c\omega_p) \lesssim 1 \tag{42}$$

in Eq. (39). Equation (42) assumes that the plasmon coupling is strongest for an ideal metal $l_c\omega_p \gg 1$. Note that the fact that g is dimensionless links the ideal metal to the large-l_c limits.

To order σ^2, the subtraction of the one-loop contribution $\mathbf{t}^{(2)}(\mathbf{k}=0,\zeta)$ and its replacement by phenomenological plasmon scattering is implemented by a (local in transverse coordinates) counterterm potential $\delta\tilde{\mathbf{V}}(\zeta,z)$ of the form

$$\delta\tilde{\mathbf{V}}^h = \mathrm{diag}(\delta V^h_{xx}(\zeta,z), \delta V^h_{yy}(\zeta,z), \left(\varepsilon\theta(z)+\frac{1}{\varepsilon}\theta(-z)\right)\delta V^h_{zz}(\zeta,z), \quad \text{with}$$
$$\delta V^h_{xx}(\zeta,z) = \delta V^h_{yy}(\zeta,z) = \delta(z)(\varepsilon-1)^2\left[\frac{-g^2\sigma^2\zeta}{1+\sqrt{\varepsilon}} + \int_0^\infty \frac{kdk}{4\pi} D(k)\left(\frac{\kappa\kappa_\varepsilon}{\varepsilon\kappa+\kappa_\varepsilon} + \frac{\zeta^2}{\kappa+\kappa_\varepsilon}\right)\right],$$
$$\delta V^h_{zz}(\zeta,z) = -\delta(z)(\varepsilon-1)^2\int_0^\infty \frac{kdk}{2\pi} D(k)\frac{k^2}{(\varepsilon\kappa+\kappa_\varepsilon)}. \tag{43}$$

Note that the support of $\delta V^h(\zeta,z)$ is in the immediate vicinity of $z=0$ only. Due to rotational and translational symmetry of the rough plate, this "counterpotential" is local and diagonal but anisotropic.[9]

As mentioned in Sec. II, the counterpotential may be interpreted as the modification of the dielectric permittivity (to order σ^2) in the vicinity of the flat interface necessary to describe the rough interface with permittivity ε and roughness correlation $D_2(\mathbf{x}-\mathbf{y})$. There is no compelling reason for perturbing about a flat interface with the *same* permittivity as the rough one. We have seen that the expansion about a flat plate with the same permittivity is not consistent with the low-energy description, since it implies unacceptably high momenta in the loop integrals. Expanding instead about the bare permittivity function of Eq. (9) yields a better controlled approximation, and Eq. (43) strongly suppresses high-momentum contributions to one loop.

IV. ROUGHNESS CORRECTION TO THE CASIMIR FREE ENERGY OF ORDER σ^2

We now evaluate the roughness correction to the Casimir free energy within the framework of the improved low-energy effective field theory. From Eqs. (23) and (27), we have altogether four contributions to order σ^2:

$$\Delta F_T^{\mathrm{Cas}}[a] = \frac{1}{2}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle - \frac{1}{2}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^|\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle + \frac{1}{2}\mathrm{Tr}\delta\tilde{\mathbf{V}}^h\mathbf{G}^{|a|} - \frac{1}{4}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle + \mathcal{O}(\sigma^3). \tag{44}$$

We consider them in turn.

A. The seagull contribution $\frac{1}{2}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle$

The first is the seagull contribution of Fig. 4(a), given by

$$\begin{aligned}\frac{1}{2}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle &= -\frac{AT}{2}\sum_n(\varepsilon-1)\left\langle\int_0^\infty\frac{kdk}{2\pi}\int_0^{h(\mathbf{x})}dz(G^{|a|}_{xx}(k,z,z;\zeta)+G^{|a|}_{yy}(k,z,z;\zeta)+\left(\varepsilon\theta(z)+\frac{\theta(-z)}{\varepsilon}\right)G^{|a|}_{zz}(k,z,z;\zeta))\right\rangle\\ &= -AT\sum_n\int_0^\infty\frac{kdk}{\pi}\kappa\kappa_\varepsilon\left(\frac{\bar{r}^2}{e^{2a\kappa}-\bar{r}^2}+\frac{r^2}{e^{2a\kappa}-r^2}\right)\left\langle\int_0^{h(\mathbf{x})}zdz\right\rangle+\mathcal{O}(\sigma^3)\\ &= -AT\sigma^2\sum_{\zeta\in\{\zeta_n\}}\int_0^\infty\frac{kdk}{2\pi}\kappa\kappa_\varepsilon\left(\frac{\bar{r}^2}{e^{2a\kappa}-\bar{r}^2}+\frac{r^2}{e^{2a\kappa}-r^2}\right)+\mathcal{O}(\sigma^3).\end{aligned} \tag{45}$$

The expressions of Eq. (A8) in Appendix VII have been expanded here for small z. There are (as expected) no corrections of order σ, and the final line exhibits equally weighted contributions from both polarizations. Note that this remarkable simplification occurs only upon summation of all δ-function contributions to $\mathbf{G}^{\parallel}_{zz}$—which gives an expansion in $\tilde{\mathbf{V}}^h$ [defined in Eq. (26)], rather than in the original $\mathbf{V}^h$.

This roughness contribution to the free energy is entirely local and does not depend on the correlation length l_c. The loop integral over transverse momenta and the sum over Matsubara frequencies are exponentially restricted to momenta $2a\kappa \lesssim 1$, and the evaluation of the seagull diagram

[9]The product of distributions in $\delta\tilde{\mathbf{V}}^h(\zeta,z) \propto \delta(z)(\varepsilon(\zeta)\theta(z) + \varepsilon^{-1}(\zeta)\theta(-z))$ here means that integration with a test function $f(z)\in L^0$ gives $\int dz(\varepsilon\theta(z)+\varepsilon^{-1}\theta(-z))\delta(z)f(z) \coloneqq \frac{1}{2}(\varepsilon\lim_{z\to 0_+}+\varepsilon^{-1}\lim_{z\to 0_-})f(z)$.

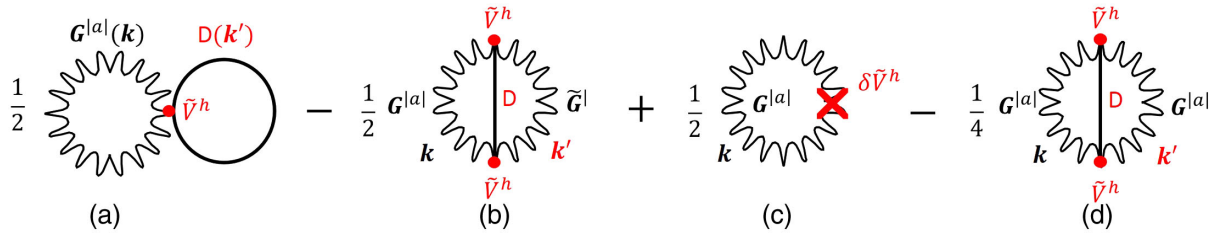

FIG. 4 (color online). Feynman diagrams for the contributions of order σ^2 to the roughness correction of the Casimir free energy of a rough and a flat interface. (a) and (b) give corrections from a single scattering off the rough surface and include only one factor of $\mathbf{G}^{|a|}$. (c) gives the contribution from the counterpotential defined in Eq. (43) whose two terms are shown in Fig. 3. This contribution eliminates the uncontrolled high-momentum contributions to the loop integral of (b) in favor of a phenomenological (plasmon) description. (d) is the 2-scattering contribution of order σ^2 and includes two factors of $\mathbf{G}^{|a|}$. The momenta in either loop of this term are exponentially restricted to $k, k' \lesssim 1/(2a) \ll \omega_p$, and no subtraction is required. Wavy lines denote photon propagators for a single flat interface, $\tilde{\mathbf{G}}^{|}(\mathbf{k}')$, or their correction, $\mathbf{G}^{|}a|(\mathbf{k})$, due to the presence of a second flat interface at a mean distance a. Solid lines represent the Fourier transform $D(\mathbf{k}-\mathbf{k}')$ of the roughness correlation function. A (red) dot indicates the effective anisotropic interaction potential $\tilde{V}^h$ due to the roughness profile defined in Eq. (26). Combinatorial factors are shown, but traces and momentum integrals have been suppressed.

using the low-energy propagators should be accurate for all $a\omega_p \gg 0.5$—that is, for $a \gtrsim 12$ nm in the case of gold plates.

Due to the κ_ε factor of the integrand, the contribution of Eq. (45) is proportional to $\omega_p\sigma^2/a^4$ for $a\omega_p \gg 1 \gg Ta$ and diverges in the ideal metal limit. Fortunately, the seagull is not the whole story to order σ^2.

B. The single diffusive scattering contribution $\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle$

The other contribution to the Casimir free energy of order σ^2 from a single scattering off the rough interface corresponds to the diagram of Fig. 4(b). This unsubtracted two-loop contribution is formally given by

$$\begin{aligned}-\frac{1}{2}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle = -\frac{AT}{2}\sum_n \int \frac{d\mathbf{k}d\mathbf{k}'}{(2\pi)^4}\,\mathrm{Tr}[&D_{++}(q)\tilde{\mathbf{G}}^{|(n)}_{++}(\mathbf{k}')\mathbf{V}^{(n)}_{+}\mathbf{G}^{|a|(n)}_{++}(\mathbf{k})\mathbf{V}^{(n)}_{+}\\ &+ D_{--}(q)\tilde{\mathbf{G}}^{|(n)}_{--}(\mathbf{k}')\mathbf{V}^{(n)}_{-}\mathbf{G}^{|a|(n)}_{--}(\mathbf{k})\mathbf{V}^{(n)}_{-} + D_{-+}(q)\tilde{\mathbf{G}}^{|(n)}_{-+}(\mathbf{k}')\mathbf{V}^{(n)}_{+}\mathbf{G}^{|a|(n)}_{+-}(\mathbf{k})\mathbf{V}^{(n)}_{-}\\ &+ D_{+-}(q)\tilde{\mathbf{G}}^{|(n)}_{+-}(\mathbf{k}')\mathbf{V}^{(n)}_{-}\mathbf{G}^{|a|(n)}_{-+}(\mathbf{k})\mathbf{V}^{(n)}_{+}],\end{aligned} \tag{46}$$

with $q = |\mathbf{k}-\mathbf{k}'|$ and the interaction vertices

$$\begin{aligned}\mathbf{V}^{(n)}_{+} &= (\varepsilon(\zeta_n)-1)\mathrm{diag}(1,1,\varepsilon(\zeta_n)),\\ \mathbf{V}^{(n)}_{-} &= (\varepsilon(\zeta_n)-1)\mathrm{diag}(1,1,1/\varepsilon(\zeta_n)).\end{aligned} \tag{47}$$

$\mathbf{G}^{(n)}_{\pm\mp}(\mathbf{k}) \coloneqq \mathbf{G}(\mathbf{k}, z=0^{\pm}, z'=0^{\mp};\zeta_n)$ denote one-sided limits of propagators. Explicit expressions are given in Eq. (A9). The correlation functions $D_{\pm\mp}(q)$ of positive and negative components of the roughness profile are defined in Eq. (30) and computed in Appendix B.

A lengthy but otherwise straightforward evaluation of Eq. (46) using the expressions of Eq. (A9) and Eq. (A3) yields

$$\begin{aligned}-\frac{1}{2}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle = -\frac{AT}{2}\sum_n(\varepsilon-1)^2\int_0^\infty\frac{kdk}{2\pi}\int_0^\infty\frac{k'dk'}{(2\pi)^2}\int_{-\pi}^{\pi}d\theta D\left(\sqrt{k^2+k'^2-2kk'\cos\theta}\right)\\ \times\left[\frac{r(1-r^2)\zeta^2}{2(e^{2a\kappa}-r^2)\kappa_\varepsilon}\left(\frac{\kappa'\kappa'_\varepsilon\sin^2\theta}{\varepsilon\kappa'+\kappa'_\varepsilon}+\frac{\zeta^2\cos^2\theta}{\kappa'+\kappa'_\varepsilon}\right)\right.\\ \left.+\frac{\bar{r}(1-\bar{r}^2)}{2(e^{2a\kappa}-\bar{r}^2)\varepsilon}\left(\frac{\varepsilon k^2k'^2-\kappa_\varepsilon^2\kappa'\kappa'_\varepsilon\cos^2\theta}{\kappa_\varepsilon(\varepsilon\kappa'+\kappa'_\varepsilon)}-kk'\bar{r}'\cos\theta-\frac{\kappa_\varepsilon\zeta^2\sin^2\theta}{(\kappa'+\kappa'_\varepsilon)}\right)\right].\end{aligned} \tag{48}$$

The signed correlation functions in Eq. (46) combine, and Eq. (48) depends on the roughness correlation $D(|\mathbf{k}-\mathbf{k}'|)$ only. In Appendix C, the integral over θ in Eq. (48) is performed analytically for the class of correlations $D_s(q)$, but this angular integral in general has to be evaluated numerically. More importantly, the leading term of order ω_p in the limit $\omega_p \to \infty$ of Eq. (48) cancels the leading asymptotic behavior $\propto \omega_p$ of the seagull term in Eq. (45).

The limit of Eq. (48) for large correlation lengths $l_c \gg 1/\omega_p$ is found using Eq. (40) to trivially evaluate the $\mathbf{k}'$ integrals. Some algebraic manipulations simplify the expression in this limit to

$$-\frac{1}{2}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|a|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle \overset{l_c\to\infty}{\longrightarrow} AT\sigma^2\sum_n\int_0^\infty\frac{kdk}{2\pi}\kappa(\kappa_\varepsilon-\kappa)\times\left(\frac{r^2}{e^{2a\kappa}-r^2}+\frac{\bar{r}^2}{e^{2a\kappa}-\bar{r}^2}\right). \quad (49)$$

C. The counterterm correction

As for $\mathbf{t}^{(2)}$ in Eq. (29), the loop integral of Eq. (48) generally includes high-momentum contributions $k' \gg \omega_p$ for which the low-energy description is not justified. The same one-loop counterpotential of Eq. (43) that corrects roughness corrections to the scattering matrix to one loop also removes the uncontrolled high-momentum contributions to the Casimir free energy and replaces them with the phenomenological plasmon contribution.

The correction of the Casimir free energy by this counterpotential is shown diagrammatically in Fig. 4(c), and the two Feynman diagrams of this counterterm are depicted in Fig. 3. To order σ^2, the contribution to the Casimir free energy from the counterpotential $\delta\tilde{\mathbf{V}}$ of Eq. (43) is

$$\frac{1}{2}\mathrm{Tr}\delta\tilde{\mathbf{V}}\mathbf{G}^{|a|}=\frac{AT}{2}\sum_n(\varepsilon-1)^2\int_0^\infty\frac{kdk}{2\pi}\int_0^\infty\frac{k'dk'}{2\pi}D(k')\left[\frac{\bar{r}(1-\bar{r}^2)k^2}{2(e^{2a\kappa}-\bar{r}^2)\kappa_\varepsilon}\frac{k'^2}{(\varepsilon\kappa'+\kappa'_\varepsilon)}+\left(\frac{r(1-r^2)\zeta^2}{2(e^{2a\kappa}-r^2)\kappa_\varepsilon}-\frac{\bar{r}(1-\bar{r}^2)\kappa_\varepsilon}{2(e^{2a\kappa}-\bar{r}^2)\varepsilon}\right)\left(\frac{\kappa'\kappa'_\varepsilon/2}{\varepsilon\kappa'+\kappa'_\varepsilon}+\frac{\zeta^2/2}{\kappa'+\kappa'_\varepsilon}-\frac{g^2\zeta}{1+\sqrt{\varepsilon}}\right)\right]. \quad (50)$$

This correction to the Casimir free energy remains finite in the ideal metal limit when Eq. (41) is satisfied. The existence of this limit is assured by the consistency of the low-energy theory in the limit $l_c \gg 1/\omega_p$. Using Eq. (40), the counterterm correction of Eq. (50) for $l_c \gg \omega_p$ becomes

$$\frac{1}{2}\mathrm{Tr}\delta\tilde{\mathbf{V}}\mathbf{G}^{|a|}\overset{l_c\to\infty}{\longrightarrow}AT\sigma^2\sum_n(g^2-1)\zeta(\sqrt{\varepsilon}-1)\int_0^\infty\frac{kdk}{2\pi}\kappa\left(\frac{r^2}{e^{2a\kappa}-r^2}+\frac{\bar{r}^2\kappa_\varepsilon^2}{(e^{2a\kappa}-\bar{r}^2)(\varepsilon k^2+\kappa_\varepsilon^2)}\right)\overset{g^2\to1}{\longrightarrow}0 \quad (51)$$

and vanishes when Eq. (41) is enforced. This should be expected of a model that is valid at low energies. Note that the reason magnetic and electric modes do not enter the counterterm correction symmetrically even at large correlation lengths is that we subtracted at $\mathbf{k}=0$: the factor $\kappa_\varepsilon^2/(\varepsilon k^2+\kappa_\varepsilon^2)$ in Eq. (51) differs from unity in order $k^2/{\omega_p}^2$ only.

D. Contributions of second order in the roughness scattering matrix

Both loop integrals of this contribution [represented in Fig. 4(d)] to the Casimir free energy are exponentially constrained to low momenta k, $k' \lesssim 1/(2a) \ll \omega_p$—a regime in which the low-energy description is expected to hold. We find that

$$-\frac{1}{4}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|a|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle=-\frac{AT}{16}\sum_n(\varepsilon-1)^2\int_0^\infty\frac{kdk}{2\pi}\int_0^\infty\frac{k'dk'}{(2\pi)^2}\int_{-\pi}^{\pi}d\theta D\left(\sqrt{k^2+k'^2-2kk'\cos\theta}\right)\times\left[\frac{r(1-r^2)\zeta^2}{(e^{2a\kappa}-r^2)\kappa_\varepsilon}\left(\frac{r'(1-r'^2)\zeta^2\cos^2\theta}{(e^{2a\kappa'}-r'^2)\kappa'_\varepsilon}-\frac{2\bar{r}'(1-\bar{r}'^2)\kappa'_\varepsilon\sin^2\theta}{(e^{2a\kappa'}-\bar{r}'^2)\varepsilon}\right)+\frac{\bar{r}\bar{r}'(1-\bar{r}^2)(1-\bar{r}'^2)}{(e^{2a\kappa}-\bar{r}^2)(e^{2a\kappa'}-\bar{r}'^2)}\left(\frac{k^2k'^2}{\kappa_\varepsilon\kappa'_\varepsilon}+\frac{2kk'\cos\theta}{\varepsilon}+\frac{\kappa_\varepsilon\kappa'_\varepsilon\cos^2\theta}{\varepsilon^2}\right)\right]. \quad (52)$$

For profiles with large correlation lengths $l_c \gg 2a \gtrsim 1/\omega_p$, Eq. (52) simplifies to

$$-\frac{1}{4}\langle \mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|a|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle\overset{l_c\to\infty}{\longrightarrow}-AT\sigma^2\sum_n\int_0^\infty\frac{kdk}{2\pi}\kappa^2\left(\frac{r^4}{(e^{2a\kappa}-r^2)^2}+\frac{\bar{r}'^4}{(e^{2a\kappa'}-\bar{r}'^2)^2}\right), \quad (53)$$

when Eq. (40) holds.

E. The limit $l_c \gg \max(1/\omega_p, a)$: the Proximity Force Approximation

Although $l_c \gg a$ is a necessary condition for the PFA, the limiting expressions of Eqs. (49) and (51) evidently hold only when l_c is large compared to *a and* $1/\omega_p$. The latter restriction arises because the scattering matrix locally can be approximated by a flat surface only if the plasma length is shorter than the typical length scale of the surface structure.

For a rough profile with $l_c \gg \max(1/\omega_p, a)$, Eqs. (49), (51), and (53) should all be reasonable approximations. Including the seagull term of Eq. (45), the roughness correction to the Casimir free energy of Eq. (44) in the limit of large correlation length $l_c \gg \max(1/\omega_p, a)$ is

$$\Delta F_T^{\mathrm{Cas}}[a] \overset{l_c\to\infty}{\longrightarrow} -AT\sigma^2\sum_n\int_0^\infty\frac{kdk}{2\pi}\kappa^2\left(\frac{r^4}{(e^{2a\kappa}-r^2)^2}+\frac{r^2}{e^{2a\kappa}-r^2}+\frac{\bar{r}'^4}{(e^{2a\kappa'}-\bar{r}'^2)^2}+\frac{\bar{r}^2}{e^{2a\kappa}-\bar{r}^2}\right)$$
$$=\frac{1}{2}\sigma^2\frac{\partial^2}{\partial a^2}\frac{AT}{2}\sum_n\int_0^\infty\frac{kdk}{2\pi}\ln(1-r^2e^{-2a\kappa})+\ln(1-\bar{r}^2e^{-2a\kappa})=\frac{1}{2}\sigma^2\frac{\partial^2}{\partial a^2}F_T^{\parallel}(a), \tag{54}$$

where $F_T^{\parallel}(a)$ is the Casimir free energy for two flat, parallel semi-infinite slabs at a separation a given by Eq. (12). This is precisely the roughness correction in PFA for a rough surface with $\langle h(\mathbf{x})\rangle = 0$ and $\langle h^2(\mathbf{x})\rangle = \sigma^2$. Although it is trivial, one should note that the PFA here emerges in the limit of large l_c from requiring consistency of the low-energy effective field theory. It is due to the absence of high-momentum contributions in this limit and does not require any additional phenomenological input.

F. Ideal metal limit $\varepsilon \to \infty$

It perhaps is remarkable that the requirement of Eq. (41) not only guarantees that the PFA is recovered in the $l_c \to \infty$ limit, but it also ensures the existence of an ideal metal limit. If g^2 is analytic at $\zeta = 0$, one can argue that ζ/ω_p and $1/(l_c\omega_p)$ [see Eq. (58)] corrections are absent and g^2 for large ω_p has the expansion $g^2 = 1 + \mathcal{O}(\zeta^2/{\omega_p}^2)$. The ideal metal limit in this case is uniquely given by

$$\frac{1}{2}\langle\mathrm{Tr}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle-\frac{1}{2}\langle\mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle=-AT\sum_n\int_0^\infty\frac{kdk}{2\pi}\int_0^\infty\frac{k'dk'}{(2\pi)^2}\int_{-\pi}^{\pi}d\theta D\left(\sqrt{k^2+k'^2-2kk'\cos\theta}\right)$$
$$\times\left[\frac{(\zeta^2+kk'\cos\theta)^2+\kappa^2\kappa'^2}{\kappa\kappa'(e^{2a\kappa}-1)}-\frac{4k\zeta^2(k-k'\cos\theta)e^{2\kappa a}}{\kappa^2(e^{2a\kappa}-1)^2}\right], \tag{55a}$$

$$-\frac{1}{4}\langle\mathrm{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|a|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle=-AT\sum_n\int_0^\infty\frac{kdk}{2\pi}\int_0^\infty\frac{k'dk'}{(2\pi)^2}\int_{-\pi}^{\pi}d\theta D\left(\sqrt{k^2+k'^2-2kk'\cos\theta}\right)$$
$$\times\frac{(\zeta^2+kk'\cos\theta)^2+\kappa^2\kappa'^2}{(e^{2a\kappa}-1)(e^{2a\kappa'}-1)\kappa\kappa'}, \tag{55b}$$

$$\frac{1}{2}\mathrm{Tr}\delta\tilde{\mathbf{V}}\mathbf{G}^{|a|}=AT\sum_n\int_0^\infty\frac{kdk}{2\pi}\int_0^\infty\frac{k'dk'}{2\pi}D(k')\left[\frac{2k^2k'^2+(\kappa^2+\zeta^2)(\kappa'-\zeta)^2}{2(e^{2a\kappa}-1)\kappa\kappa'}\right]. \tag{55c}$$

Note that the counterterm contribution of Eq. (55c) does not vanish and cancels the contribution from high-k' momenta in Eq. (55a) also for the ideal metal. High-momentum contributions to the roughness correction thus persist in the ideal metal limit. Without the counterterm, this perturbative correction would diverge for $l_c \to 0$ (and for some correlations would diverge for *all* l_c). This apparently is at odds with exact calculations for square-wave profiles [16] and demands an explanation. The reason for convergence of the exact calculations in the limit $l_c \to 0$ (and divergence of the unsubtracted perturbation theory) for such profiles is subtle and related to the fact that for $l_c \ll \sigma$, the leading term in the exact calculation is $\mathcal{O}(\sigma)$ and not $\mathcal{O}(\sigma^2)$ as perturbation theory suggests [16]. The nonanalytic dependence on σ for $l_c \to 0$ arises due to an effective UV cutoff in the exact calculation of $\mathcal{O}(\sigma)$—there is no other scale to compare with in this limit. Ignoring this effective cutoff (as a perturbative expansion in σ does) leads to an UV-divergent expression in the limit $l_c \to 0$. The nonanalyticity of the exact result for $\sigma/a \ll 1$ in the limit $l_c \to 0$ is only possible if wave numbers of order $1/\sigma$ of the profile contribute significantly. The nonanalyticity in σ in this sense implies that high-momenta $1/a < k' < 1/\sigma$ must dominate the exact Casimir energy calculation for an ideal metal in the limit $0 \le l_c < \sigma \sim 0$. A simple model that qualitatively reproduces this explanation of the nonanalytic dependence on σ is obtained by replacing $l_c \to l_c + \gamma\sigma$ in the Gaussian correlation function of Eq. (34), where the constant γ is of $\mathcal{O}(1)$. For $l_c \gg \sigma$, one recovers the quadratic perturbative dependence on σ in leading order, but for $0 \le l_c \ll \sigma \to 0$, the k' integral of Eq. (55a) is of order $\sigma^2/(l_c+\gamma\sigma) \overset{l_c\ll\sigma}{\to} \sigma/\gamma$, as in the exact calculation. The UV divergence $\propto 1/\sigma^3$ of the k' integral that gives this leading (nonanalytic) behavior is due to momenta $k' \sim 1/\sigma \gg 1/a$. Although the exact evaluation of such high-momentum contributions is of itself correct, the low-energy description

used to compute them is not justified. The fact that the plasmon contributes—and the counterterm of Eq. (55c) removes—high-momentum contributions even for an ideal metal indirectly supports the assertion that roughness corrections of real materials in fact remain analytic in the variance σ^2 even in the limit of uncorrelated roughness.

G. The limit of uncorrelated roughness and the plasmon coupling g^2

The high-roughness limit $l_c \ll 1/\omega_p$ is obtained by examining the loop integrals in Eqs. (45), (48), and (50) at large momentum transfers $q = |\mathbf{k}' - \mathbf{k}|$. In the limit of uncorrelated roughness $l_c \to 0$, the correction is

$$\Delta F_T^{\mathrm{Cas}}[a] \overset{l_c\to 0}{\to} -AT\sigma^2 \sum_n \int_0^\infty \frac{kdk}{2\pi} \left[\frac{\bar{r}^2 \kappa \kappa_\varepsilon (2\frac{\varepsilon-1}{\varepsilon+1}k^2 + \kappa_\varepsilon^2 - g^2(\sqrt{\varepsilon}-1)\zeta\kappa_\varepsilon)}{(e^{2a\kappa} - \bar{r}^2)(k^2\varepsilon + \kappa_\varepsilon^2)} + \frac{r^2\kappa(\kappa_\varepsilon - g^2(\sqrt{\varepsilon}-1)\zeta)}{e^{2a\kappa} - r^2} \right]. \tag{56}$$

Note that the correction to the Casimir free energy for $l_c = 0$ is strictly negative when $g^2 \leq 1$. The Casimir free energy of a rough interface thus is always larger in magnitude than that of a flat one at the same average separation. We believe this is the result of two opposing effects. The specular reflection off a rough surface with vanishing l_c but finite σ never is quite the same as that off a flat interface with the *same* bulk permittivity: the situation is analogous to the change in bulk permittivity due to the inclusion of subwavelength spheres of a different material. Since the included "material" in this case is vacuum, with $\varepsilon = 1$, the effective reflection coefficient decreases compared to that for the flat plate. This effect by itself would tend to decrease the Casimir free energy in magnitude for $l_c \to 0$. However, this decrease is more than compensated for by the reduced separation to this effective interface.

The ideal metal limit of Eq. (56) exists only for $g^2 \to 1$ and is analytically given by

$$\Delta F_T^{\mathrm{Cas}}[a, l_c \ll 1/\omega_p \to 0] = -AT\sigma^2 \sum_n \int_0^\infty \frac{kdk}{2\pi} \frac{\zeta(\kappa^2 + \zeta^2)}{\kappa(e^{2a\kappa} - 1)} \overset{T\to 0}{\longrightarrow} -\frac{9A\sigma^2}{32\pi^2 a^5}\zeta(5) \approx -0.02955 \frac{A\sigma^2}{a^5}. \tag{57}$$

The ideal metal and $l_c \to 0$ limits, in fact, commute, and $g^2 \to 1$ is required for the ideal metal limit to exist. Assuming that $g^2(\zeta l_c, l_c\omega_p)$ is analytic in both arguments, the existence of an ideal metal limit implies

$$1 = \lim_{\substack{\omega_p\to\infty \\ l_c\omega_p = \beta}} g^2(\zeta l_c, l_c\omega_p) = g^2(0, \beta). \tag{58}$$

We therefore have that $g^2 = 1$ at low frequencies for any value of l_c and ω_p. We in the following, therefore, consider only

$$g^2 = 1. \tag{59}$$

V. THE EFFECTIVE LOW-ENERGY FIELD THEORY OF ELECTROMAGNETIC INTERACTIONS WITH ROUGH SURFACES

Although we obtained a roughness correction that is compatible with the low-energy theory of Schwinger by a Green's function approach, it is instructive to construct the effective low-energy field theory from which these corrections derive. The effective field theory allows one to, in principle, explore other approximations and corrections. It also provides a general framework for systematically taking into account higher orders or for including other interactions. In this formulation, the necessity of counterterms, furthermore, is readily apparent.

A. The generating functional of roughness correlations

The construction of the field theory is based on the generating function of the roughness correlation functions rather than the roughness correlations themselves. This approach was already used in the scalar case [22]. The n-point roughness correlation functions for an interface of (large) area A with a *particular* profile $h(\mathbf{x})$ are the averages

$$\begin{aligned} D_1 &= \langle h(\mathbf{x}_1)\rangle &&:= A^{-1}\int_A h(\mathbf{x}+\mathbf{x}_1)d\mathbf{x}, \\ D_2(\mathbf{x}_1-\mathbf{x}_2) &= \langle h(\mathbf{x}_1)h(\mathbf{x}_2)\rangle &&:= A^{-1}\int_A h(\mathbf{x}+\mathbf{x}_1)h(\mathbf{x}+\mathbf{x}_2)d\mathbf{x}, \\ &\vdots &&\vdots \\ D_n(\mathbf{x}_1-\mathbf{x}_2,\ldots,\mathbf{x}_{n-1}-\mathbf{x}_n) &= \langle h(\mathbf{x}_1)\ldots h(\mathbf{x}_n)\rangle &&:= A^{-1}\int_A h(\mathbf{x}+\mathbf{x}_1)\ldots h(\mathbf{x}+\mathbf{x}_n)d\mathbf{x}. \end{aligned} \tag{60}$$

The interface is assumed to be large enough for boundary effects to be negligible. Transverse translational invariance then implies that these correlations depend only on *differences* of the transverse coordinates.[10] Isotropy of the roughness profile yields further restrictions: the n-point correlation function in this case depends only on *distances* between the points. We assume that the profile, and therefore all n-point correlation functions of Eq. (60), can, at least in principle, be measured when the rough interface is far removed from the other. The mean separation a between the two interfaces is determined so that Eq. (4) holds—that is, the (constant) one-point function D_1 vanishes. We formally collect all roughness correlation functions of Eq. (60) for a particular profile $h(\mathbf{x})$ in a single generating functional $Z_h[\alpha]$:

$$Z_h[\alpha] = \sum_{n=2}^{\infty} \frac{1}{n!} \iint \alpha(\mathbf{x}_1)\alpha(\mathbf{x}_2)\ldots\alpha(\mathbf{x}_n) \times D_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)d\mathbf{x}_1 d\mathbf{x}_2\ldots d\mathbf{x}_n. \tag{61}$$

Note that Z_h depends on a *particular* profile $h(\mathbf{x})$. None of the D_n are averages over different profiles. For another profile, some or all of the correlations defined by Eq. (60) change, and so does the functional $Z_h[\alpha]$.

However, for constructing the field theory, it is expedient to directly model $Z_h[\alpha]$ instead of computing the individual correlation functions of a given profile $h(\mathbf{x})$. With the restriction of Eq. (4) that the one-point function vanishes, the simplest model of a rough interface is entirely determined by its two-point correlation D_2. The generating functional of such a (quadratic) Gaussian model is of the form

$$Z_h^G[\alpha] = \exp\frac{1}{2}(\alpha|D_2|\alpha), \tag{62}$$

with

$$(\alpha|D_2|\alpha) \coloneqq \iint \alpha(\mathbf{x})D_2(\mathbf{x}-\mathbf{y})\alpha(\mathbf{y})d\mathbf{x}d\mathbf{y}. \tag{63}$$

In general, Eq. (62) is only the leading quadratic term in a cumulant expansion of Z_h. A Gaussian generating functional relates all higher-order correlations to the two-point function. In Appendix B, for instance, we determine signed correlation functions in terms of D_2 for such a model. Stochastic roughness is fully described by the covariance of the profile, and a Gaussian model by definition is exact in this case. A Gaussian model for the generating functional also suffices to obtain corrections to the free energy and the scattering matrix to leading order in the variance of the roughness profile. To order σ^2, the correlations of a periodic one-dimensional profile $h^\omega(\mathbf{x}) = \sigma\sin(\omega x)$ can be found using a Gaussian model, but the four-point correlation function in this case is only half of what the Gaussian model asserts,

$$D_2^\omega(\mathbf{x}-\mathbf{y}) = \frac{\sigma^2}{2}\cos(\omega(x-y)), \qquad \text{but}$$
$$D_4^\omega(\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3,\mathbf{x}_4) = \frac{1}{2}(D_2^\omega(\mathbf{x}_1-\mathbf{x}_2)D_2^\omega(\mathbf{x}_3-\mathbf{x}_4) + D_2^\omega(\mathbf{x}_1-\mathbf{x}_3)D_2^\omega(\mathbf{x}_2-\mathbf{x}_4) + D_2^\omega(\mathbf{x}_1-\mathbf{x}_4)D_2^\omega(\mathbf{x}_2-\mathbf{x}_3)). \tag{64}$$

To correctly obtain effects due to a periodic profile to order σ^4 requires the inclusion of a fourth-order cumulant. Note that the two-point correlation D_2^ω in Eq. (64) of a periodic corrugated profile is not positive definite and has no probabilistic interpretation. However, in momentum space, it is proportional to the sum of two δ functions and is therefore positive semidefinite.

The basis for a field-theoretic approach to roughness is that *any* analytic functional $R[h]$ of the profile $h(\mathbf{x})$ with *translation-invariant* coefficients can be evaluated using $Z_h[\alpha]$. To show this, consider a typical monomial in the Taylor expansion of $R[h]$ for small profiles:

$$\begin{aligned}
&\iint d\mathbf{x}_1 d\mathbf{x}_2\ldots d\mathbf{x}_n R_n(\mathbf{x}_1-\mathbf{x}_2,\ldots,\mathbf{x}_{n-1}-\mathbf{x}_n)h(\mathbf{x}_1)h(\mathbf{x}_2)\ldots h(\mathbf{x}_n)\\
&= \frac{1}{A}\int_A d\mathbf{x} \iint d\mathbf{x}_1 d\mathbf{x}_2\ldots d\mathbf{x}_n R_n(\mathbf{x}_1-\mathbf{x}_2,\ldots,\mathbf{x}_{n-1}-\mathbf{x}_n)h(\mathbf{x}+\mathbf{x}_1)h(\mathbf{x}+\mathbf{x}_2)\ldots h(\mathbf{x}+\mathbf{x}_n)\\
&= \iint d\mathbf{x}_1 d\mathbf{x}_2\ldots d\mathbf{x}_n R_n(\mathbf{x}_1-\mathbf{x}_2,\ldots,\mathbf{x}_{n-1}-\mathbf{x}_n)D_n(\mathbf{x}_1-\mathbf{x}_2,\ldots,\mathbf{x}_{n-1}-\mathbf{x}_n)\\
&= \iint d\mathbf{x}_1 d\mathbf{x}_2\ldots d\mathbf{x}_n R_n(\mathbf{x}_1-\mathbf{x}_2,\ldots,\mathbf{x}_{n-1}-\mathbf{x}_n)\frac{\delta}{\delta\alpha(\mathbf{x}_1)}\frac{\delta}{\delta\alpha(\mathbf{x}_2)}\ldots\frac{\delta}{\delta\alpha(\mathbf{x}_n)}Z_h[\alpha]\Big|_{\alpha=0}.
\end{aligned} \tag{65}$$

[10]For exact translational invariance, the finite, parallel flat surfaces could be replaced by two concentric two-dimensional tori of area A.

The first equality in Eq. (65) is due to the translational invariance of the coefficient functions R_n [but it does not assume any regularity of the profile $h(\mathbf{x})$ itself]. No further assumptions are required to show that Eq. (65) holds for *any* profile of a sufficiently large interface. The second equality in Eq. (65) implies that the result is proportional to the area A. Assuming that all coefficient functions R_n in the Taylor expansion of $R[h]$ are translation invariant and that the expansion converges for the particular profile, Eq. (65) implies that one may evaluate $R[h]$ for any *particular* profile $h(\mathbf{x})$ by applying the corresponding functional derivative operator on $Z_h[\alpha]$,

$$R[h] = R\left[\frac{\delta}{\delta\alpha}\right] Z_h[\alpha]|_{\alpha=0}. \tag{66}$$

B. The partition function of the low-energy effective field theory

In the presence of external sources of polarization $\vec{P}_n(\mathbf{x}, z) = \vec{P}(\mathbf{x}, z; \zeta_n)$, Schwinger's free energy for two parallel interfaces is given by Eq. (11). The partition function for a flat and a rough interface described by the profile $h(\mathbf{x})$ corresponding to the potential $\mathbf{V}(\zeta, h(\mathbf{x}), z)$ of Eq. (19), therefore, formally is

$$\begin{aligned} Z_T[\vec{P}, h] = &\exp\left[-\frac{1}{T}(F_T^{\parallel}(a) + \delta F[h])\right] \\ &\times \prod_n \exp\left[-\frac{1}{T}(\mathcal{V}_n[h] + \delta\mathcal{V}_n^h)\right] \\ &\times \exp\left[\frac{T}{2}\{\vec{P}_n|\mathbf{G}^{\parallel(n)}|\vec{P}_n\}\right], \end{aligned} \tag{67}$$

where $\mathcal{V}_n[h]$ is the functional derivative operator,

$$\mathcal{V}_n[h] = -\frac{1}{2}\int d\mathbf{x} \int_0^{h(\mathbf{x})} dz \frac{\delta}{\delta\vec{P}_n(\mathbf{x},z)} \cdot (\boldsymbol{\varepsilon}(\zeta_n) - \mathbb{1}) \cdot \frac{\delta}{\delta\vec{P}_n^{\dagger}(\mathbf{x},z)}, \tag{68}$$

representing the interaction of the nth Matsubara mode with the roughness profile $h(\mathbf{x})$.

C. Counterterms of the low-energy effective field theory

The counterpotential of Eq. (43) corresponds to a functional derivative operator of the form

$$\delta\mathcal{V}_n^h = \frac{1}{2}\int dz \int d\mathbf{x} \frac{\delta}{\delta\vec{P}_n(\mathbf{x},z)} \cdot \delta\mathbf{V}^h(\zeta_n, z) \cdot \int d\mathbf{y} \frac{\delta}{\delta\vec{P}_n^{\dagger}(\mathbf{y},z)}. \tag{69}$$

It corrects for polarization effects due to surface roughness. Note that the counterpotential of Eq. (43) in Eq. (69) has support in the immediate vicinity of the plane at $z = 0$ only and does not depend on the transverse position $\mathbf{x}$ or on the mean separation a of the two interfaces. The counterpotential $\delta\mathbf{V}^h(\zeta)$ should ensure that the scattering of electromagnetic waves incident perpendicular to the rough surface is reproduced.

We in addition have to include a counterterm $\delta F[h]$ to the free energy that is a functional of the profile $h(\mathbf{x})$. It vanishes for $h(\mathbf{x}) = 0$ and has the expansion

$$\begin{aligned} \delta F[h] = c_0 &+ \int d\mathbf{x} h(\mathbf{x}) c_1(a,T) \\ &+ \frac{1}{2}\iint d\mathbf{x} d\mathbf{y} c_2(\mathbf{x}-\mathbf{y}) h(\mathbf{x}) h(\mathbf{y}) \\ &+ \frac{1}{6}\iiint d\mathbf{x} d\mathbf{y} dz c_3(\mathbf{x}-z, \mathbf{y}-z) h(\mathbf{x}) h(\mathbf{y}) h(z) \\ &+ \cdots, \end{aligned} \tag{70}$$

with translation-invariant n-point coefficient functions c_n that depend only on transverse coordinate differences. These coefficient functions are used to systematically remove corrections to the correlation functions of the profile $h(x)$ in the presence of electromagnetic interactions. The constant one-point counterterm $c_1(a,T)$ ensures that $\langle h(x)\rangle = 0$ at any separation a and temperature T. $c_1(a,T)$ is the only coefficient that may depend on a and T, because its contribution to the free energy in fact vanishes for profiles that satisfy Eq. (4). The higher-order terms of $\delta F[h]$ are constructed so that connected correlation functions of the profile at $T = 0$ are the prescribed ones when the second flat interface is removed. They do not depend on the temperature T or on the separation a. This ensures that

$$\begin{aligned} &\frac{\partial}{\partial T}\delta F[h] = \frac{\partial}{\partial a}\delta F[h] = 0 \\ &\quad \text{for any profile for which } \int_A d\mathbf{x} h(\mathbf{x}) = 0. \end{aligned} \tag{71}$$

This counterterm to the free energy, therefore, does not affect thermodynamic state functions like the entropy or pressure. It cancels loop contributions to the energy (at $T \to 0$) when the flat interface is removed ($a \to \infty$). The Casimir free energy remains (its finite, a-dependent value at $T = 0$ is the Casimir energy).

In obtaining the Casimir free energy by the Green's function method, the contribution to the free energy from the counterterm coefficient $c_2(\mathbf{x}-\mathbf{y})$ was implicitly taken into account by subtracting $\Delta F_T[h, \infty]$ in Eq. (21). Requiring the absence of one-loop corrections to the two-point roughness correlation at large separation a and temperature $T = 0$ determines $c_2(\mathbf{q})$. The Feynman diagrams involved in this condition are shown in Fig. 5. The counterterm c_2 also ensures that there is no single-interface correction to the Casimir energy at $T = 0$. For $T > 0$, a

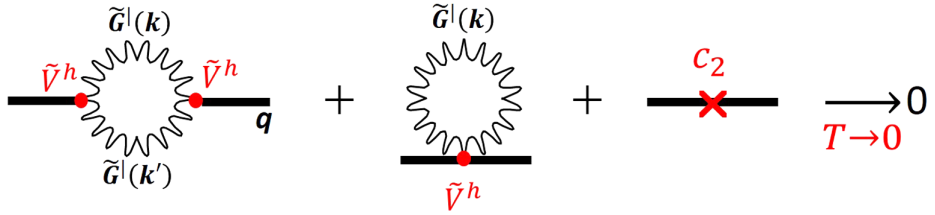

FIG. 5 (color online). One-loop Feynman diagrams for the counterterm $c_2(q)$. $c_2(q)$ is determined by demanding that the (prescribed) two-point roughness correlation of a single plate at $T=0$ is not corrected. We here consider one-loop contributions only.

finite a-independent contribution to the single-interface free energy remains that we have not calculated here.

The Green's function approach implicitly also accounts for contributions of $c_1(a,T)$ by simply assuming that Eq. (4) holds to order σ^2. $c_1(a,T)$ cancels tadpole contributions to the scattering matrix (see Fig. 6), and one-particle reducible contributions to the Casimir free energy like those of Fig. 7 vanish in this case.

We define the mean separation a by Eq. (4), and demanding that corrections to $\langle h_\pm(\mathbf{x})\rangle$ vanish determines $c_1(a,T)$ to one loop. The diagrammatic form of this condition is shown in Fig. 6 and evaluates to

$$\begin{aligned} c_1(a,T) &= \frac{T}{D(0)}\sum_n \int \frac{d\mathbf{k}}{(2\pi)^2}\left[D_{\pm+}(0)\mathrm{Tr}\mathbf{V}_+^{(n)}(\tilde{\mathbf{G}}_{++}^{|(n)}(\mathbf{k})+\mathbf{G}_{++}^{|a|(n)}(\mathbf{k}))+D_{\pm-}(0)\mathrm{Tr}\mathbf{V}_-^{(n)}(\tilde{\mathbf{G}}_{--}^{|(n)}+\mathbf{G}_{--}^{|a|(n)}(\mathbf{k}))\right] \\ &= c_1(\infty,T)-T\sum_n\int_0^\infty \frac{kdk}{2\pi}\kappa\left(\frac{\bar{r}^2}{e^{2\kappa a}-\bar{r}^2}+\frac{\bar{r}^2}{e^{2\kappa a}-\bar{r}^2}\right) \\ &= c_1(\infty,T)-\frac{\partial}{A\partial a}F_T^{\parallel}(a), \end{aligned} \tag{72}$$

where $c_1(\infty,T)$ is the (infinite) one-interface contribution that does not depend on the separation a. The interpretation of Eq. (72) is straightforward and could have been anticipated: for $\langle h\rangle\neq 0$, the separation a is redefined at one loop, since

$$F_T^{\parallel}(a)-\int_A d\mathbf{x}h(\mathbf{x})\frac{\partial}{A\partial a}F_T^{\parallel}(a)\approx F_T^{\parallel}(a-\langle h\rangle). \tag{73}$$

To leading order in $\langle h\rangle$, the c_1 counterterm arises from the free energy of two parallel flat interfaces at separation a_B, where $a=a_B+\langle h\rangle$ is the separation at which Eq. (4) holds.

The a-independent but temperature-dependent contribution from $c_1(\infty,T)$, similarly, is the difference in free energy due to a shift of a flat interface by $-\langle h\rangle$. The bulk contribution to the free energy density thereby increases by

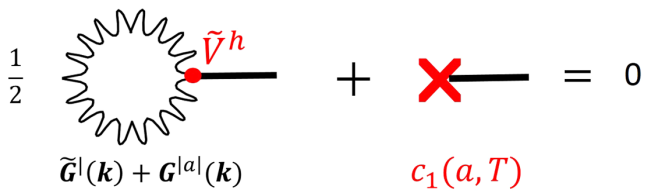

FIG. 6 (color online). Cancellation of tadpoles by the counterterm $c_1(a,T)$ at one loop. Summation to all orders of the δ-function contribution to $G_{zz}^{\parallel}$ replaces $\mathbf{G}^{\parallel}$ with $\tilde{\mathbf{G}}^{|}+\mathbf{G}^{|a|}$ and $\mathbf{V}^h$ by $\tilde{\mathbf{V}}^h$.

$$\begin{aligned} c_1(\infty,T) &= -\frac{T}{4}\sum_n\int_0^\infty\frac{kdk}{2\pi}\mathrm{Tr}(V_+^{(n)}\tilde{\mathbf{G}}_{++}^{|(n)}(k)+V_-^{(n)}\tilde{\mathbf{G}}_{--}^{|(n)}(k)) \\ &= -\frac{T}{2}\sum_n(\varepsilon(\zeta_n)-1)\int_0^\infty\frac{kdk}{2\pi}\left(\frac{\kappa_\varepsilon\kappa-k^2}{\varepsilon\kappa+\kappa_\varepsilon}+\frac{\zeta^2}{\kappa_\varepsilon+\kappa}\right) \\ &= T\sum_n\int_0^\infty\frac{kdk}{2\pi}(\kappa-\kappa_\varepsilon)=\frac{1}{V}(F_T^\gamma[1]-F_T^\gamma[\varepsilon]), \end{aligned} \tag{74}$$

where $F_T^\gamma[\varepsilon]/V$ is the free energy density of a photon gas in a homogeneous medium with permittivity $\varepsilon(\zeta)$. The difference in free energy density in the dielectric and in vacuum depends on the permittivity $\varepsilon(\zeta)$. For the plasma model with $\varepsilon(\zeta)=1+(\omega_p/\zeta)^2$, this separation-independent contribution to the free energy is

$$\begin{aligned} (F_T^\gamma[1]-F_T^\gamma[\varepsilon])\frac{A\langle h\rangle}{V} = A\langle h\rangle\Bigg[c_1(\infty,0)-\frac{T^4\pi^2}{45} \\ +\frac{T^2{\omega_p}^2}{\pi^2}\sum_{n=1}^\infty\frac{K_2(n\omega_p/T)}{n^2}\Bigg], \end{aligned} \tag{75}$$

where the modified Bessel function $K_2(x)$ is normalized to $K_2(x\sim 0)\sim 2/x^2$. The generally infinite constant $c_1(\infty,0)$ does not depend on temperature or on the separation a. It is sensitive to the behavior of $\varepsilon(\zeta)$ at energies $\zeta\gg\omega_p$. Estimating this contribution to the free energy in the framework of the low-energy effective theory is

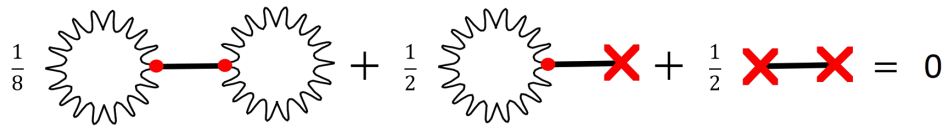

FIG. 7 (color online). One-particle reducible dumbbell contributions to the free energy that are canceled by the c_1 counterterm given in Eq. (72). One-particle reducible contributions to the free energy are of order $1/T$ at low temperatures and would violate Nernst's theorem.

meaningless, since the loop integral is dominated by momenta and energies $k, \zeta \gg \omega_p$. For the sake of completeness, this formal contribution with a proper time cutoff β is

$$c_1(a \sim \infty, T = 0) = -\frac{1}{16\pi^2}\int_\beta^\infty \frac{d\lambda}{\lambda^3}(1 - e^{-\lambda\omega_p{}^2}). \quad (76)$$

It is a quadratically and logarithmically UV-divergent constant contribution to the total energy of the system. It may be absorbed in the counterterm c_0 and in the absence of gravitational interactions has no physical implications.

D. The complete low-energy effective field theory

Since the Green's function $\mathbf{G}^\parallel$ of parallel interfaces, as well as the counterterms, are invariant under transverse translations, the partition function $Z_T(\vec{P} = 0, h)$ defined in Eq. (67) for vanishing polarization sources is a functional of the roughness profile $h(\mathbf{x})$ with translation-invariant coefficients. We thus can use Eq. (66) to evaluate it using the correlation functions of the profile $h(\mathbf{x})$ rather than the profile itself. We therefore have that

$$Z_T[\vec{P} = 0, h] = Z_T\left[\vec{P} = 0, \frac{\delta}{\delta\alpha}\right] Z_h[\alpha]|_{\alpha=0}, \quad (77)$$

with $Z_\alpha[h]$ defined by Eq. (61). The complete generating functional of the Gaussian model we are considering thus is

$$\begin{aligned}\mathcal{Z}_T[\vec{P}, \alpha] \coloneqq & \exp\left[-\frac{1}{T}(F_T^\parallel(a) + \delta F[\delta/\delta\alpha])\right] \\ & \times \prod_n \exp\left[-\frac{1}{T}(\mathcal{V}_n[\delta/\delta\alpha] + \delta\mathcal{V}_n^h)\right] \\ & \times \exp\left[\frac{T}{2}\{\vec{P}_n|\mathbf{G}^{\parallel(n)}|\vec{P}_n\} + \frac{1}{2}(\alpha|D_2|\alpha)\right], \quad (78)\end{aligned}$$

with $(\alpha|D_2|\alpha)$ given by Eq. (63). The partition function of Eq. (77) is just $\mathcal{Z}_T[\vec{P} = 0, \alpha = 0]$. From the point of view of Euclidean field theory, Eq. (78) promotes the roughness profile $h(\mathbf{x})$ to a field on the two-dimensional (planar) subspace that is coupled to a vector field in $\mathbb{R}^3 \times S_1$. Correlation functions of $h(\mathbf{x})$ are obtained by functional differentiation of Eq. (78) with respect to the scalar source α, and $\mathcal{Z}$ defines the loop expansion in the usual manner. The main difference from ordinary field theory is that all correlation functions of $h(\mathbf{x})$ are prescribed, and counterfunctions enforce the absence of any corrections to them at $T = 0$ and $a \sim \infty$. The low-energy effective field theory encoded by Eq. (78) evidently is not renormalizable—new counterterms (functions) are required at each order of the loop expansion. The three counterterms c_1, c_2 of $\delta F[h]$ and $\delta\mathbf{V}^h$ suffice at the one-loop level, since only the connected two-point functions and $\langle h\rangle$ are superficially UV dominated if $D_2(0) = \sigma^2$ is finite.

Instead of employing the Green's function approach, one can derive the loop corrections to the free energy from Eq. (78). The Casimir free energy to one loop is the same in both approaches. However, the generating functional Eq. (78) of the low-energy effective theory has conceptual and methodical advantages: once the set of counterterms is determined, the field theory yields consistent low-energy results not just for the Casimir energy, but for the scattering matrix as well. No *ad hoc* arguments or procedures are required to cancel uncontrolled high-energy loop corrections, and the necessity of the counterterms and their interpretation is readily apparent.

VI. NUMERICAL INVESTIGATIONS

We numerically investigated the correction $\Delta F_T^{\mathrm{Cas}}(a)$ to the Casimir free energy given in Eq. (44) due to the roughness of an interface. To order σ^2, this correction is linear in the roughness correlation function, and one may define [17] a response function $R_T(q, a)$,

$$\Delta F_T^{\mathrm{Cas}}(a) = \int_0^\infty \frac{qdq}{2\pi} R_T(q, a) D(q), \quad (79)$$

that does not depend on $D(q)$. Analytical expressions for $R_T(q, a)$ are obtained by changing the integration variable from $\mathbf{k}'$ to $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ in Eqs. (45), (48), (50), and (52). The corresponding expressions are given in Appendix D. For clarity, and to compare with earlier investigations, we in the following present numerical results for $T = 0$ only. Temperature corrections are sizable only when $2\pi aT \gtrsim 1$. For gold surfaces at 300 °K, temperature corrections become important at separations of the order of microns—a distance at which perturbative roughness is irrelevant.

A. The response with and without the counterterm

Figure 8 gives the normalized response when the counterterm of Eq. (50) is omitted as a function of the dimensionless variable q/ω_p. The low-energy theory is in the shaded momentum region $q/\omega_p < 1$. Note the linear rise of the low-energy response function for all separations a in the uncontrolled region $q/\omega_p \gg 1$. The integration weight $qD(q)$ for Gaussian and exponential roughness

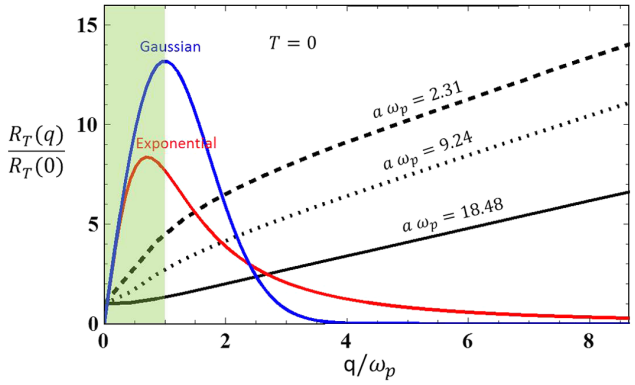

FIG. 8 (color online). The dimensionless normalized response $\rho(q,a) = R_T(q,a)/R_T(0,a)$ without counterpotential $\delta\mathbf{V}^h = 0$ for the permittivity $\varepsilon(\zeta) = 1 + (\omega_p/\zeta)^2$ to leading order in σ^2 at $T = 0$. The dependence on q/ω_p of this ratio of the roughness response function $R_T(q,a)$ [defined by Eq. (79)] is shown for $a\omega_p = 2.31$ (−−), 9.24 (· · · ·), and 18.48 (——). For the plasma frequency $\omega_p = \omega_p(\mathrm{Au}) \sim 0.046\ \mathrm{nm}^{-1}$, this normalized response without counterpotential is identical to that obtained by Ref. [17]. (For $\omega_p = 0.046\ \mathrm{nm}^{-1}$, the curves here correspond to those of Fig. 4 in Ref. [17] at separations $a = 50$, 100, and 200 nm.) Note the change in behavior and the subsequent linear rise in the region $q\omega_p \gtrsim 1$. The region $q\omega_p \lesssim 1$, where the effective low-energy theory is valid, is shaded light green. We superimpose typical integration densities for the response function in Eq. (79): the momentum-space function $qD(q)$ for Gaussian and exponential two-point roughness correlation with $l_c = 1/\omega_p$. The roughness correction to the Casimir energy with exponential correlation diverges logarithmically, and even for Gaussian roughness correlation, the (unshaded) region $q/\omega_p > 1$ contributes significantly in this uncorrected case. Note that for a gold surface, the correlation length here is $l_c = 1/\omega_p(\mathrm{Au}) \sim 21$ nm.

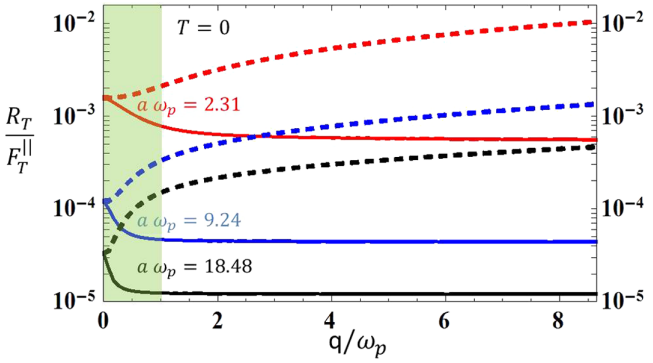

FIG. 9 (color online). The ratio $R_T(q,a)/F_T^{\parallel}(a)$ of the roughness response function to the Casimir energy of flat parallel plates at $T = 0$ with (solid) and without (dashed) counterpotential $\delta\mathbf{V}^h$ with $g^2 = 1$. The permittivity $\varepsilon(\zeta) = 1 + (\omega_p/\zeta)^2$ is characterized by the plasma frequency ω_p. The dependence on q/ω_p of the ratio is shown for $a\omega_p = 2.31$ (top, red), 9.24 (middle, blue), and 18.48 (bottom, black). For $\omega_p = 0.046\ \mathrm{nm}^{-1} \sim \omega_p(\mathrm{Au})$, the normalized response without counterpotential (dashed) is identical to that of Fig. 3 in Ref. [17] at separations of $a = 50$, 100, and 200 nm. Note that the renormalized roughness response is monotonically decreasing and approaches a constant at large momenta that are a factor of 2–3 smaller than the response at $q = 0$. Most of the correction to the Casimir energy in this case arises from the shaded integration region $q/\omega_p < 1$, where the low-energy description is valid.

correlation with a typical correlation length $l_c \sim 1/\omega_p$ is superimposed. A sizable contribution to the roughness correction in Eq. (79) evidently is due to loop momenta $q > \omega_p$, for which low-energy expressions are unreliable.

Inclusion of the counterpotential gives a constant high-momentum response. Figure 9 shows the response functions with and without the counterterm contribution of Eq. (50). With the same model for the bulk permittivity of gold, the response function shown in Fig. 3 of Ref. [17] is reproduced when the counterpotential is omitted. Inclusion of the counterpotential gives a constant high-momentum response, and the correction to the Casimir (free) energy is of order σ^2. Note that with $g^2 = 1$, the response at $q = 0$ does not change.

The correction to the Casimir energy at $T = 0$ for Gaussian roughness with and without inclusion of the counterterm of Eq. (50) is shown in Fig. 10. Whereas the PFA limit $l_c \to \infty$ coincides for both cases, the behavior is remarkably different at finite l_c. Including the counterterm of Eq. (50), the roughness correction to the Casimir energy *decreases* in magnitude for decreasing correlation length and approaches a finite (uncorrelated) limit for $l_c \to 0$. Roughness increases the Casimir force, but the PFA is an upper bound in this case. The ratio of the roughness correction to the PFA furthermore approaches a constant, l_c-dependent value with increasing separation rather than increasing indefinitely as in the unsubtracted case [for exponential roughness, the roughness correction without the counterterm of Eq. (50) would diverge at any separation and for all l_c]. Let us also note that for $l_c \lesssim 1/\omega_p$, the roughness correction at large separations is less than 50% of the PFA prediction. Although we here are considering only perturbative roughness corrections, the suppression at large separations for $l_c \lesssim 1/\omega_p$ is of a similar magnitude to that observed [25] for machined profiles with correlation length $l_c \sim 1/\omega_p$.

B. (In)sensitivity on high-momentum components of the roughness correlation

The counterpotential $\delta\mathbf{V}^h$ was introduced to correct for uncontrolled high-momentum contributions to loop integrals with the help of phenomenological input. We therefore investigated the sensitivity of the roughness correction to the correlation function $D(q)$ numerically. Figure 11 shows the ratio of the correction for Gaussian and for exponential roughness of the same correlation length l_c. The two are identical for $l_c = 0$ and $l_c \sim \infty$ (PFA) at any $a\omega_p$. The (dimensionless) ratio of these corrections never drops below 85% for any separation $a\omega_p$ and correlation length $l_c\omega_p$. Without counterpotential, this ratio is infinite. Exponential roughness always gives a smaller correction than Gaussian

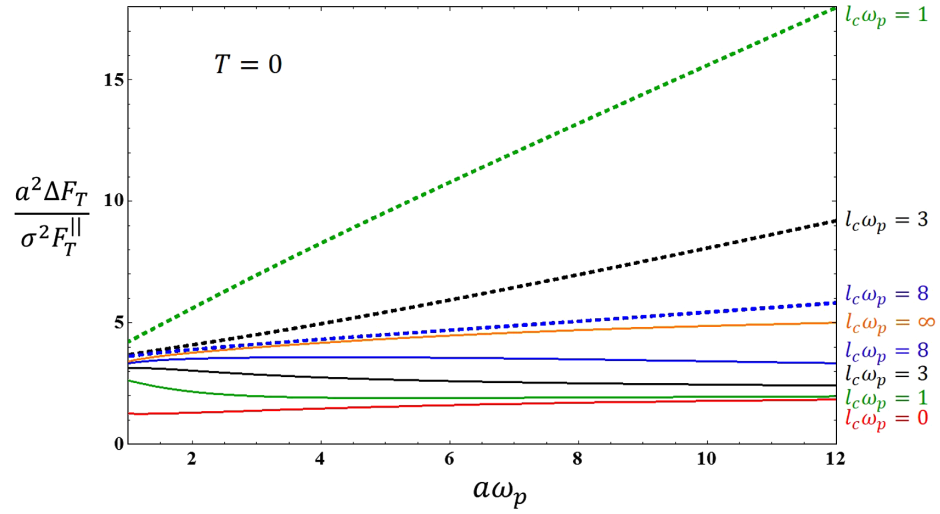

FIG. 10 (color online). The dimensionless ratio $(a^2/\sigma^2)\Delta F_T^{\mathrm{Cas}}(a)/F_T^{\parallel}(a)$ of the roughness correction to the Casimir energy of two parallel flat interfaces at $T = 0$. The calculation is to leading order in σ^2/a^2 for a plasma-model permittivity with plasma frequency ω_p for Gaussian roughness with correlation length l_c. Dashed curves give the ratio as a function of $a\omega_p$ without the counterterm contribution of Eq. (50), whereas solid curves give the ratio when this counterterm with $g^2 = 1$ is included. Curves of the same color correspond to the same value of $l_c\omega_p$. From the top: $l_c\omega_p = 1$ (green, dashed), 3 (black, dashed), 8 (blue, dashed), ∞ (orange), 8 (blue, solid), 3 (black, solid), 1 (green, solid), and 0 (red, solid). Note that the $l_c \to 0$ curve (red) is a lower bound that exists only in the renormalized case. The counterterm vanishes in the PFA limit $l_c \to \infty$ (orange), and this limit is the same for both. Whereas the PFA is an upper bound for the magnitude of the roughness correction when the counterpotential is included, it is a lower bound without. The ratio of the roughness correction to the PFA at finite l_c approaches a finite value at large separations when the counterterm is included, whereas it otherwise increases indefinitely. The roughness correction in the subtracted case at large separations is less than 50% of the PFA prediction when $l_c \lesssim 1/\omega_p$. Except for $l_c = 0$, the roughness correction approaches the PFA estimate at sufficiently small separation, but it quickly decreases and approaches the lower bound for $l_c\omega_p < 1$.

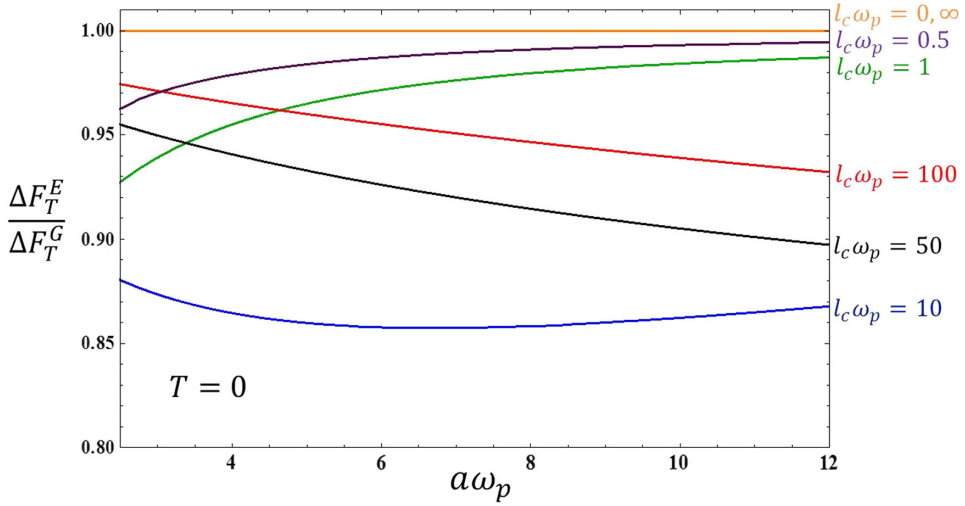

FIG. 11 (color online). The dimensionless ratio $\Delta F_T^E(a)/\Delta F_T^G(a)$ of the roughness correction to the Casimir energy for exponential (E) and Gaussian (G) roughness with the same correlation length $l_c\omega_p$ as a function of the dimensionless separation $a\omega_p$. $g^2 = 1$, and a plasma-model permittivity characterized by the single plasma frequency ω_p was assumed. The roughness correlation functions are those of Eq. (35) (E) and Eq. (34) (G). In the PFA ($l_c \to \infty$) and uncorrelated ($l_c \to 0$) limits, the corrections coincide but differ by up to 15% at some separations. For the same variance σ^2 and correlation length l_c, the roughness correction with exponential correlation is always smaller than with Gaussian correlation. Note that the two types of roughness correlation approach the PFA quite differently: at large separations the corrections still differ by over 5% even for $l_c\omega_p \sim 100$.

roughness of the same correlation length and variance. The two correlation functions provide rather similar descriptions of low-energy scattering, and the low-energy effective theory with counterpotential depends only weakly on their (very different) behavior at high momenta.

C. Comparison with experiment

The low-energy theory for electromagnetic interactions with rough surfaces ultimately must be compared to experiment. Unfortunately, very few studies are dedicated to the systematic investigation of Casimir forces between rough surfaces. Many employ nonisotropic machined surfaces with rather large σ/a ratios [24,25] that are not accessible perturbatively. Nevertheless, these experiments qualitatively contradict the predictions of exact calculations, that essentially any kind of roughness tends to increase the Casimir force above the PFA estimate. A notable exception is a series of investigations of isotropically rough surfaces by Palasantzas *et al.* [11,12]. For sufficiently rough surfaces, this group does observe (see Fig. 3 of Ref. [12]) an increase of the Casimir force by 200%–400% at small separations. This sharp increase in the force was attributed to particularly high islands of the surface profile that can also be seen in some of the AFM scans of the gold surfaces. The pronounced effect of such islands is beyond the scope of a perturbative analysis and was explained by a semiempirical approach [13] based on the PFA.

However, gold films with 100 nm and 200 nm thickness and relatively low roughness appear to be almost free of such buildup effects. At small separations, the force in these cases is smaller than the PFA prediction. In Fig. 12, we

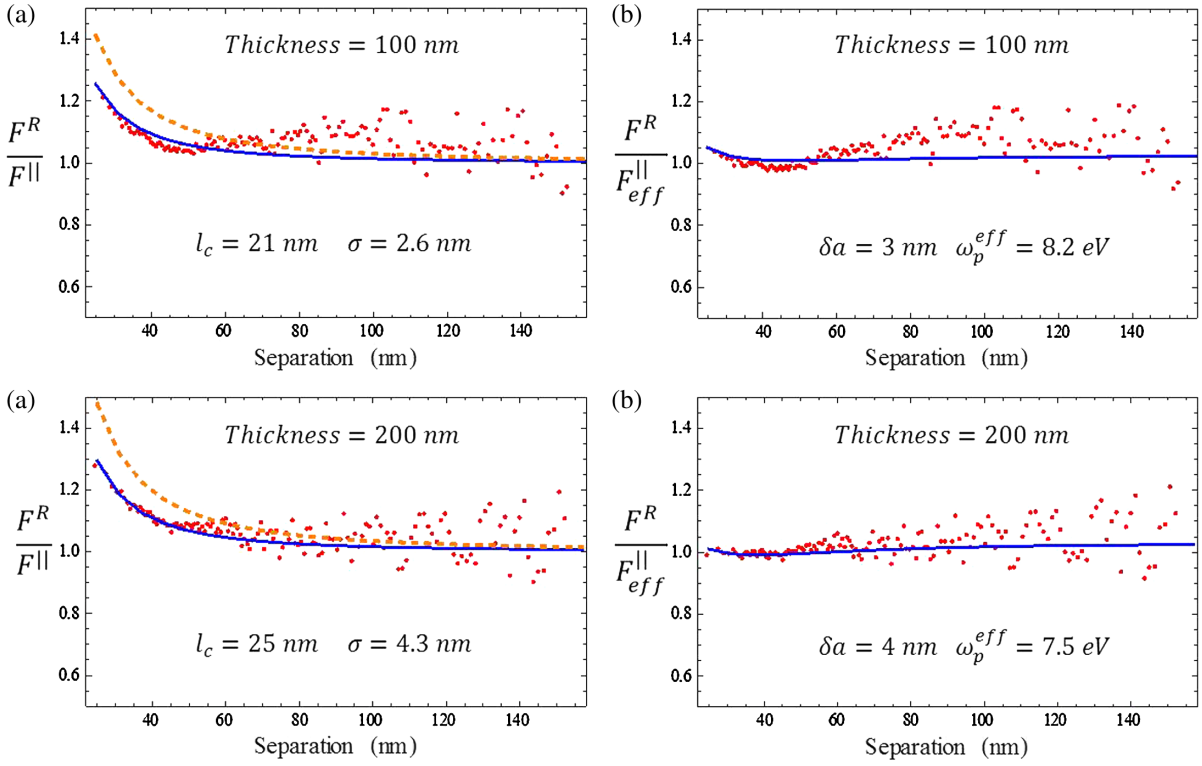

FIG. 12 (color online). The dimensionless ratio $\rho(a)$ defined in Eq. (82) of the Casimir force between a rough gold-coated sphere and a rough gold-coated plate to the Casimir energy between ideal dielectric flat plates. The experimental data is from Ref. [12]. The thickness of the gold coating on the flat plate is 100 nm (upper graphs) and 200 nm (lower graphs). An exponential roughness correlation and a Drude parametrization of the permittivity is assumed. The standard deviation and correlation length for the sphere's profile are $\sigma^{\mathrm{Sph}} \sim 8$ nm and $l_c^{\mathrm{Sph}} \sim 33$ nm. (a) The ratio of the force on the rough plate to the Casimir force between a gold-coated flat plate and a smooth sphere at the same mean separation. A Drude parametrization of the permittivity with $\omega_p = 9$ eV, $\gamma = 0.045$ eV was used. (Red) dots denote the ratio for the experimental data of Ref. [12]. The measured force on the 100 nm thick plate was multiplied by a correction factor of 0.94 (see text for details). The solid (blue) line is our best theoretical fit to this ratio with the indicated parameters for the roughness correlation function of the plate in Eq. (80). Note that the ~30% enhancement at separations $a \sim 20$ nm is well reproduced for both films. The dashed line gives the PFA for roughness of the same total variance. (b) The ratio of the force on the rough plate to that between a smooth sphere and a flat plate at the separation $a - \delta a$. The indicated $\omega_{p\mathrm{eff}}$ for the effective permittivity of the flat plate was obtained from ellipsometric measurements [12] on the rough ones. We assumed the same effective plasma frequency $\omega_{p\mathrm{eff}}^{\mathrm{Sph}} = 7.5$ eV for the sphere as for the (similarly rough) 200 nm film. The solid (blue) line gives the ratio to the force on the effective flat plate and sphere for the same force including the roughness corrections shown in (a). Note that this ratio of the force with roughness corrections to that between a flat plate and smooth sphere with the measured reflection coefficients at a reduced separation is close to unity for all separations.

compare the low-energy theory to the measurements of Ref. [12] on these thin films. The experiments measure the force between a gold-coated sphere and a gold-coated plate. Both surfaces are rough, but their profiles are uncorrelated. For two parallel rough gold-coated plates, the correction to the Casimir energy to leading order in σ/a is that for a single rough plate with a roughness correlation that is the sum of the roughness correlations functions of the sphere and the flat plate,

$$D(q) = D^{\text{plate}}(q) + D^{\text{sphere}}(q). \tag{80}$$

We use Derjaguin's PFA approximation [5] to correct for the curvature of the sphere of radius $R = 100\ \mu\text{m} \gg a$. The force $f_T(a)$ at temperature T between the sphere and a plate with (closest) separation a in this approximation is

$$f_T(a) = 2\pi R F_T^{\text{Cas}}[a]/A, \tag{81}$$

where $F_T^{\text{Cas}}[a]/A$ is the Casimir free energy per unit area (not the pressure) of two parallel rough plates. Due to the large radius of the sphere, this is an excellent approximation for separations $a < 200\ \text{nm} \sim R/500$. Fig. 12(a) gives the ratio $\rho(a)$ of this force to the Casimir energy per unit area $F_T^{\parallel}/A$ of two *flat* parallel gold plates with separation a,

$$\rho(a) \coloneqq \frac{f_T(a)A}{2\pi R F_T^{\parallel}} = \frac{F_T^{\text{Cas}}[a]}{F_T^{\parallel}} = 1 + \frac{\Delta F_T^{\text{Cas}}[a]}{F_T^{\parallel}}, \tag{82}$$

at $T = 0$. The experimental Casimir force for the rough sphere and plate at separations $\sigma \ll a < l_c$ is up to 30% greater than the Casimir energy for flat plates.

Since we do not differentiate between contributions from high and low peaks of the roughness profile and only use a single correlation function, all standard deviations of Ref. [12] were multiplied by a factor of 1.7. We used $\sigma^{\text{Sph}} = 8$ nm, $\sigma^{100} = 2.6$ nm, and $\sigma^{200} = 4.3$ nm for the coatings of the sphere, 100 nm and 200 nm thick films, respectively. These standard deviations also approximately correspond to those estimated from the AFM scans of these surfaces (see Fig. 1 in Ref. [12]). The correlation lengths $l_c^{\text{Sph}} = 33$ nm, $l_c^{100} = 21$ nm, and $l_c^{200} = 25$ nm are those of Ref. [12]. The ratio $\rho(a)$ for the 200 nm thick film is well reproduced by the low-energy theory with exponential as well as with Gaussian correlations. We only show the result for exponential roughness in Fig. 12, but the fit for Gaussian roughness is of similar quality. For comparison, we show the roughness correction in PFA for the same standard deviations.

The ratio $\rho(a)$ is close to unity at larger separations $100\ \text{nm} < a < 150$ nm, where roughness corrections are relatively small. While this on average is approximately observed for the 200 nm film, the ratio for the 100 nm film is systematically about 6% above unity at larger distances. To correct for this (unexplained) discrepancy, we multiplied the force observed on the 100 nm thick film by 0.94 before[11] comparing with theory.

From a practical point of view, the comparison in Fig. 12(b) with the Casimir energy of two parallel flat plates at a slightly smaller separation $a_{\text{eff}} = a - \delta a$ perhaps is more useful. The Drude model permittivity describing reflection off these effective flat plates in Ref. [12] was obtained from ellipsometric measurements on the rough surfaces. We merely adjusted δa for the best fit. Figure 12(b) shows that effective flat surfaces at a reduced separation $a - \delta a$ reproduce the low-roughness data remarkably well. [The force data of the 100 nm film was multiplied by the same correction factor of 0.94 as in the graph of Fig. 12(a).] Since ellipsometric measurements on thin films are quite standard, this observation essentially reduces low-roughness corrections to Casimir energies to a determination of the optimal shift δa. Instead of measuring the absolute average distance between the profiles of two rough surfaces (in itself a delicate procedure that involves a number of corrections), we suggest that precision Casimir studies with low-roughness surfaces simply determine an *effective* separation for flat plates with the measured (perpendicular) reflection coefficients. Figure 12(b) is evidence that the data at small separations robustly determines this distance to better than 1 nm, at the same time all but eliminating the need for roughness corrections.

VII. CONCLUSION

We obtained roughness corrections to low-energy scattering and the Casimir free energy in the framework of Schwinger's effective theory of low-energy electrodynamics. The energy scale in this theory is the plasma frequency $\omega_p \sim 0.046\ \text{nm}^{-1} \sim 9$ eV of typical materials like gold. We found that roughness corrections generally include large contributions from high-momentum excitations. Evaluating them in the low-energy framework is inconsistent and notoriously unreliable. We emphasize that this is not a limitation of the perturbative approach developed here: exact (numerical) solutions of a model can also only be as accurate as the model itself. The Casimir energy of short-wavelength periodic rectangular profiles, for instance, involves momenta at which a description in terms of the bulk permittivity of the material breaks down, and the mathematically exact analysis of such a model can lead to physically erroneous conclusions. Using the bulk permittivity to describe scattering off profile structures with sizes of the order of the inverse plasma frequency or smaller (about 25 nm for gold) is not justified. Effects due to

[11] While this correction factor is *ad hoc*, we would like to point out that the ratios of Fig. 12 are less forgiving than logarithmic depictions of the data. The experimental error probably increases sharply at larger separations simply because the force is rapidly decreasing in magnitude.

roughness on the scale of the plasma frequency generally are grossly overestimated by the uncorrected low-energy theory. This has been experimentally verified for machined profiles with a period $\lambda \lesssim 2\pi/\omega_p$: the exact calculations [38,39] for such profiles tend to overestimate the observed [25] Casimir force by factors of 2–3.

We presented a perturbative analysis of roughness corrections based on a low-energy effective field theory that employs counterterms to correct for uncontrolled high-momentum contributions. The counterterms subtract high-momentum contributions to loop integrals at the cost of phenomenological input. Apart from correlations of the roughness profile itself, we in addition modeled the averaged single-interface scattering matrix at vanishing transverse momentum by the plasmon contribution. To leading order in the roughness variance σ^2, this semi-empirical *Ansatz* depends on a single coupling constant g^2. Consistency of the low-energy theory and the existence of an ideal metal limit at any correlation length constrains this dimensionless coupling to $g^2 = 1$ at low energies [see Eq. (59)]. The resulting low-energy theory is free of high-momentum contributions to one-loop integrals, approaches the PFA for $l_c \sim \infty$ and has a finite ideal metal limit for any l_c. It is relatively insensitive to the high-momentum behavior of the roughness correlation function and has a drastically different but more transparent dependence on l_c than the uncorrected model. Instead of large (infinite) differences, roughness correlation functions that differ only at high momenta now give similar low-energy predictions. Roughness of shorter correlation length no longer increases the Casimir force (indefinitely). Instead, the magnitude of the force decreases with decreasing correlation length and approaches a finite lower bound for uncorrelated roughness.

Although the coupling g^2 in the plasmon contribution to the counterterm potential Eq. (43) was constrained to $g^2 = 1$ by self-consistency and the existence of certain limits of the effective low-energy theory, this is a *model* for the roughness contribution to the average scattering matrix at low transverse momenta. It may be phenomenologically preferable to parameterize empirical data for this component of the scattering matrix instead. However, there is some evidence that the plasmon describes low-energy scattering due to roughness reasonably well. In this sense, it is a reasonable model for the leading roughness correction that is relatively simple and consistent with the low-energy theory.

Interestingly, the PFA is accurate at small separations only for $l_c \gtrsim 1/\omega_p$, and at large separations it may overestimate the correction to the force by up to 250% (see Fig. 10). For $l_c \lesssim 1/\omega_p$, the roughness correction to the Casimir energy is significantly (a factor $\sim 1/2$–$1/3$) below the PFA prediction at all but the smallest separations. The ratio remains approximately constant for $a \sim \infty$ and does not increase with increasing separation as in the uncorrected model. Although we considered only isotropic roughness profiles, it perhaps is interesting that the reduction of the correction compared to the PFA prediction by a factor of 2 for $l_c \sim 1/\omega_p$ is of the same order of magnitude as the experimental reduction in the overall force observed [25] by experiments with corrugated rectangular wave profiles.

The Casimir energy of low-roughness profiles was found to be essentially that of flat plates with the *measured* reflection coefficients at a distance that is slightly smaller than the mean separation of the interfaces. The change in separation is less than the standard deviation of the rough profile. Although the precise value of this shift depends on properties of the profile, this observation enables one to empirically correct for (low-level) roughness and accurately calibrate the effective separation in the plate-sphere geometry.

For conceptual reasons, we here derived all expressions for the Casimir free energy at finite temperature, but we only investigated implications of this theory at $T = 0$. We intend to extend the numerical investigations to finite temperature in the future. Although the roughness correction at finite temperature is not expected to change at small separations, the regime $1 < a/l_c < aT$, where temperature and roughness corrections are of similar importance, could be of some interest. At this point, we only wish to observe that the summands in all expressions at finite temperature are finite when $\zeta \to 0$ for any reasonable permittivity function (Drude or plasma model). Predictions of this low-energy effective field theory at temperatures $2\pi T > \omega_p \sim 2 \times 10^4\,^\circ\mathrm{K}$ nevertheless would be meaningless.

ACKNOWLEDGMENTS

We would like to thank G. Palasantzas for giving us access to the experimental data of his group. Discussions with K. V. Shajesh and Junming Liu provided insights that are gratefully acknowledged. H.-Y. W. enjoyed the support and hospitality of the Lorentz Center in Leiden, Netherlands, and the invitation to PASI2012, where some preliminary results were presented. This work was supported by NSF Grant No. PHY-09-02054.

APPENDIX A: THE GREEN'S DYADIC FOR THREE FLAT DIELECTRIC SLABS

In Schwinger's formalism [3], the parallel-plate Green's dyadic is determined by reduced electric and magnetic Green's functions. In the coordinate system in which $\mathbf{k} = (k, 0)$ points along the $+x$ axis, this Green's dyadic is

$$\mathbf{G}^{\parallel}(k,z,z';\zeta,a)=\begin{bmatrix} -\frac{1}{\varepsilon_z}\frac{\partial}{\partial z}\frac{1}{\varepsilon_{z'}}\frac{\partial}{\partial z'}g_H & 0 & -\frac{ik}{\varepsilon_z\varepsilon_{z'}}\frac{\partial}{\partial z}g_H \\ 0 & \zeta^2 g_E & 0 \\ \frac{ik}{\varepsilon_z\varepsilon_{z'}}\frac{\partial}{\partial z'}g_H & 0 & \frac{1}{\varepsilon_z}\delta(z-z')-\frac{k^2}{\varepsilon_z\varepsilon_{z'}}g_H \end{bmatrix}, \tag{A1}$$

where g_E and g_H solve the differential equations

$$\begin{aligned}\left[-\frac{\partial^2}{\partial z^2}+k^2+\zeta^2\varepsilon_z\right]g_E(k,z,z';\zeta)&=\delta(z-z'),\\ \left[-\frac{\partial}{\partial z}\frac{1}{\varepsilon_z}\frac{\partial}{\partial z}+\frac{k^2}{\varepsilon_z}+\zeta^2\right]g_H(k,z,z';\zeta)&=\delta(z-z').\end{aligned} \tag{A2}$$

One recovers the Green's function for arbitrary transverse momentum $\mathbf{k}$ by rotation about the z axis:

$$\begin{aligned}\mathbf{G}^{\parallel}(\mathbf{k},z,z';\zeta,a)&=\mathbf{R}\cdot\mathbf{G}^{\parallel}(k=|\mathbf{k}|,z,z';\zeta,a)\cdot\mathbf{R}^T,\\ \mathbf{R}&=\frac{1}{k}\begin{pmatrix}k_x & -k_y & 0\\ k_y & k_x & 0\\ 0 & 0 & k\end{pmatrix}.\end{aligned} \tag{A3}$$

The solution to Eq. (A2) in different regions of z and z' will be denoted

$$g_i(k,z,z';\zeta)=\begin{bmatrix} g_i^{++}(k,z>0,z'>0;\zeta) & g_i^{+-}(k,z>0,z'<0;\zeta)\\ g_i^{-+}(k,z<0,z'>0;\zeta) & g_i^{--}(k,z<0,z'<0;\zeta)\end{bmatrix}\quad\text{with}\quad i=E\quad\text{or}\quad H. \tag{A4}$$

We divide the reduced Green's functions into $g_i^{|}$ for a single flat plate and its correction $g_i^{|a|}$ due to the presence of a parallel flat plate at a distance a:

$$g_i(k,z,z';\zeta,a)=g_i^{|}(k,z,z';\zeta)+g_i^{|a|}(k,z,z';\zeta,a), \tag{A5}$$

$$\begin{aligned}g_E^{|}(k,z,z';\zeta)&=\begin{bmatrix}\frac{1}{2\kappa_2}(e^{-\kappa_2|z-z'|}-r_2e^{-\kappa_2(z+z')}) & \frac{1}{\kappa_2+\kappa_3}e^{\kappa_3 z'-\kappa_2 z}\\ \frac{1}{\kappa_2+\kappa_3}e^{\kappa_3 z-\kappa_2 z'} & \frac{1}{2\kappa_3}(e^{-\kappa_3|z-z'|}+r_2e^{\kappa_3(z+z')})\end{bmatrix},\\ g_H^{|}(k,z,z';\zeta)&=\begin{bmatrix}\frac{1}{2\bar\kappa_2}(e^{-\kappa_2|z-z'|}-\bar r_2e^{-\kappa_2(z+z')}) & \frac{1}{\bar\kappa_2+\bar\kappa_3}e^{\kappa_3 z'-\kappa_2 z}\\ \frac{1}{\bar\kappa_2+\bar\kappa_3}e^{\kappa_3 z-\kappa_2 z'} & \frac{1}{2\bar\kappa_3}(e^{-\kappa_3|z-z'|}+\bar r_2e^{\kappa_3(z+z')})\end{bmatrix},\\ g_E^{|a|}(k,z,z';\zeta,a)&=\frac{r_1}{e^{2a\kappa_3}-r_1r_2}\begin{bmatrix}\frac{1}{2\kappa_2}(1-r_2^2)e^{-\kappa_2(z+z')} & \frac{1}{\kappa_2+\kappa_3}(e^{-\kappa_2 z-\kappa_3 z'}+r_2e^{-\kappa_2 z+\kappa_3 z'})\\ \frac{1}{\kappa_2+\kappa_3}(e^{-\kappa_2 z'-\kappa_3 z}+r_2e^{-\kappa_2 z'+\kappa_3 z}) & \frac{1}{2\kappa_3}(e^{-\kappa_3 z}+r_2e^{\kappa_3 z})(e^{-\kappa_3 z'}+r_2e^{\kappa_3 z'})\end{bmatrix},\\ g_H^{|a|}(k,z,z';\zeta,a)&=\frac{\bar r_1}{e^{2a\kappa_3}-\bar r_1\bar r_2}\begin{bmatrix}\frac{1}{2\bar\kappa_2}(1-\bar r_2^2)e^{-\kappa_2(z+z')} & \frac{1}{\bar\kappa_2+\bar\kappa_3}(e^{-\kappa_2 z-\kappa_3 z'}+\bar r_2e^{-\kappa_2 z+\kappa_3 z'})\\ \frac{1}{\bar\kappa_2+\bar\kappa_3}(e^{-\kappa_2 z'-\kappa_3 z}+\bar r_2e^{-\kappa_2 z'+\kappa_3 z}) & \frac{1}{2\bar\kappa_3}(e^{-\kappa_3 z}+\bar r_2e^{\kappa_3 z})(e^{-\kappa_3 z'}+\bar r_2e^{\kappa_3 z'})\end{bmatrix}.\end{aligned} \tag{A6}$$

Note that continuity of E_x, E_y, and εE_z across the flat interface implies that g_E, g_H, and $\frac{1}{\varepsilon_z}\frac{\partial}{\partial z}\frac{1}{\varepsilon_{z'}}\frac{\partial}{\partial z'}g_H$ are continuous as well. The components of Eq. (A1) in different regions' domains of z and z' are

$$\tilde{G}^{|}_{xx}(k,z,z';\zeta) = -\frac{1}{\varepsilon_z}\frac{\partial}{\partial z}\frac{1}{\varepsilon_{z'}}\frac{\partial}{\partial z'}g^{|}_H = \frac{1}{2}\begin{bmatrix} \bar{\kappa}_2(e^{-\kappa_2|z-z'|}+\bar{r}_2e^{-\kappa_2(z+z')}) & \bar{\kappa}_3(1-\bar{r}_2)e^{-\kappa_2 z+\kappa_3 z'} \\ \bar{\kappa}_3(1-\bar{r}_2)e^{-\kappa_2 z'+\kappa_3 z} & \bar{\kappa}_3(e^{-\kappa_3|z-z'|}-\bar{r}_2e^{\kappa_3(z+z')}) \end{bmatrix},$$

$$\tilde{G}^{|}_{yy}(k,z,z';\zeta) = \zeta^2 g^{|}_E = \zeta^2\begin{bmatrix} \frac{1}{2\kappa_2}(e^{-\kappa_2|z-z'|}-r_2e^{-\kappa_2(z+z')}) & \frac{1}{\kappa_2+\kappa_3}e^{\kappa_3 z'-\kappa_2 z} \\ \frac{1}{\kappa_2+\kappa_3}e^{\kappa_3 z-\kappa_2 z'} & \frac{1}{2\kappa_3}(e^{-\kappa_3|z-z'|}+r_2e^{\kappa_3(z+z')}) \end{bmatrix},$$

$$\tilde{G}^{|}_{zz}(k,z,z';\zeta) = -\frac{k^2}{\varepsilon_z\varepsilon_{z'}}g^{|}_H = -k^2\begin{bmatrix} \frac{1}{2\varepsilon_2\kappa_2}(e^{-\kappa_2|z-z'|}-\bar{r}_2e^{-\kappa_2(z+z')}) & \frac{1}{\varepsilon_3\kappa_2+\varepsilon_2\kappa_3}e^{\kappa_3 z'-\kappa_2 z} \\ \frac{1}{\varepsilon_3\kappa_2+\varepsilon_2\kappa_3}e^{\kappa_3 z-\kappa_{z2} z'} & \frac{1}{2\varepsilon_3\kappa_3}(e^{-\kappa_3|z-z'|}+\bar{r}_2e^{\kappa_3(z+z')}) \end{bmatrix},$$

$$\tilde{G}^{|}_{xz}(k,z,z';\zeta) = -\frac{ik}{\varepsilon_z\varepsilon_{z'}}\frac{\partial}{\partial z}g^{|}_H = \frac{ik}{2}\begin{bmatrix} \frac{1}{\varepsilon_2}(\mathrm{sgn}(z-z')e^{-\kappa_2|z-z'|}-\bar{r}_2e^{-\kappa_2(z+z')}) & \frac{1}{\varepsilon_3}(1-\bar{r}_2)e^{-\kappa_2 z+\kappa_3 z'} \\ -\frac{1}{\varepsilon_2}(1+\bar{r}_2)e^{-\kappa_2 z'+\kappa_3 z} & \frac{1}{\varepsilon_3}(\mathrm{sgn}(z-z')e^{-\kappa_3|z-z'|}-\bar{r}_2e^{\kappa_3(z+z')}) \end{bmatrix},$$

$$\tilde{G}^{|}_{zx}(k,z,z';\zeta) = \frac{ik}{\varepsilon_z\varepsilon_{z'}}\frac{\partial}{\partial z'}g^{|}_H = \frac{ik}{2}\begin{bmatrix} \frac{1}{\varepsilon_2}(\mathrm{sgn}(z-z')e^{-\kappa_2|z-z'|}+\bar{r}_2e^{-\kappa_2(z+z')}) & \frac{1}{\varepsilon_2}(1+\bar{r}_2)e^{-\kappa_2 z+\kappa_3 z'} \\ -\frac{1}{\varepsilon_3}(1-\bar{r}_2)e^{-\kappa_2 z'+\kappa_3 z} & \frac{1}{\varepsilon_3}(\mathrm{sgn}(z-z')e^{-\kappa_3|z-z'|}+\bar{r}_2e^{\kappa_3(z+z')}) \end{bmatrix}. \quad \text{(A7)}$$

The corresponding separation-dependent part is

$$G^{|a|}_{xx}(k,z,z';\zeta,a) = \frac{-\bar{r}_1}{2(e^{2a\kappa_3}-\bar{r}_1\bar{r}_2)}\begin{bmatrix} \bar{\kappa}_2(1-\bar{r}_2^2)e^{-\kappa_2(z+z')} & \bar{\kappa}_3(e^{-\kappa_2 z}-\bar{r}_2e^{-\kappa_2 z})(e^{-\kappa_3 z'}-\bar{r}_2e^{\kappa_3 z'}) \\ \bar{\kappa}_3(e^{-\kappa_2 z'}-\bar{r}_2e^{-\kappa_2 z'})(e^{-\kappa_3 z}-\bar{r}_2e^{\kappa_3 z}) & \bar{\kappa}_3(e^{-\kappa_3 z}-\bar{r}_2e^{\kappa_3 z})(e^{-\kappa_3 z'}-\bar{r}_2e^{\kappa_3 z'}) \end{bmatrix},$$

$$G^{|a|}_{yy}(k,z,z';\zeta,a) = \frac{\zeta^2 r_1}{e^{2a\kappa_3}-r_1r_2}\begin{bmatrix} \frac{1}{2\kappa_2}(1-r_2^2)e^{-\kappa_2(z+z')} & \frac{1}{\kappa_2+\kappa_3}(e^{-\kappa_2 z-\kappa_3 z'}+r_2e^{-\kappa_2 z+\kappa_3 z'}) \\ \frac{1}{\kappa_2+\kappa_3}(e^{-\kappa_2 z'-\kappa_3 z}+r_2e^{-\kappa_2 z'+\kappa_3 z}) & \frac{1}{2\kappa_3}(e^{-\kappa_3 z}+r_2e^{\kappa_3 z})(e^{-\kappa_3 z'}+r_2e^{\kappa_3 z'}) \end{bmatrix},$$

$$G^{|a|}_{zz}(k,z,z';\zeta,a) = \frac{-k^2\bar{r}_1}{e^{2a\kappa_3}-\bar{r}_1\bar{r}_2}\begin{bmatrix} \frac{1}{2\varepsilon_2\kappa_2}(1-\bar{r}_2^2)e^{-\kappa_2(z+z')} & \frac{1}{\varepsilon_3\kappa_2+\varepsilon_2\kappa_3}(e^{-\kappa_2 z-\kappa_3 z'}+\bar{r}_2e^{-\kappa_2 z+\kappa_3 z'}) \\ \frac{1}{\varepsilon_3\kappa_2+\varepsilon_2\kappa_3}(e^{-\kappa_2 z'-\kappa_3 z}+\bar{r}_2e^{-\kappa_2 z'+\kappa_3 z}) & \frac{1}{2\varepsilon_3\kappa_3}(e^{-\kappa_3 z}+\bar{r}_2e^{\kappa_3 z})(e^{-\kappa_3 z'}+\bar{r}_2e^{\kappa_3 z'}) \end{bmatrix},$$

$$G^{|a|}_{xz}(k,z,z';\zeta,a) = \frac{ik\bar{r}_1}{2(e^{2a\kappa_3}-\bar{r}_1\bar{r}_2)}\begin{bmatrix} \frac{1}{\varepsilon_2}(1-\bar{r}_2^2)e^{-\kappa_2(z+z')} & \frac{1}{\varepsilon_3}(e^{-\kappa_2 z}-\bar{r}_2e^{-\kappa_2 z})(e^{-\kappa_3 z'}+\bar{r}_2e^{\kappa_3 z'}) \\ \frac{1}{\varepsilon_2}(e^{-\kappa_3 z}-\bar{r}_2e^{\kappa_3 z})(e^{-\kappa_2 z'}+\bar{r}_2e^{-\kappa_2 z'}) & \frac{1}{\varepsilon_3}(e^{-\kappa_3 z}-\bar{r}_2e^{\kappa_3 z})(e^{-\kappa_3 z'}+\bar{r}_2e^{\kappa_3 z'}) \end{bmatrix},$$

$$G^{|a|}_{zx}(k,z,z';\zeta,a) = \frac{-ik\bar{r}_1}{2(e^{2a\kappa_3}-\bar{r}_1\bar{r}_2)}\begin{bmatrix} \frac{1}{\varepsilon_2}(1-\bar{r}_2^2)e^{-\kappa_2(z+z')} & \frac{1}{\varepsilon_2}(e^{-\kappa_2 z}+\bar{r}_2e^{-\kappa_2 z})(e^{-\kappa_3 z'}-\bar{r}_2e^{\kappa_3 z'}) \\ \frac{1}{\varepsilon_3}(e^{-\kappa_2 z'}-\bar{r}_2e^{-\kappa_2 z'})(e^{-\kappa_3 z}+\bar{r}_2e^{\kappa_3 z}) & \frac{1}{\varepsilon_3}(e^{-\kappa_3 z}+\bar{r}_2e^{\kappa_3 z})(e^{-\kappa_3 z'}-\bar{r}_2e^{\kappa_3 z'}) \end{bmatrix}. \quad \text{(A8)}$$

The limits of these propagators as z and z' approach 0 are of particular interest. In this case, the components of the matrices $\tilde{\mathbf{G}}^{|}(k;\zeta) \coloneqq \tilde{\mathbf{G}}^{|}(k,0,0;\zeta)$ and $\mathbf{G}^{|a|}(k;\zeta,a) \coloneqq \mathbf{G}^{|a|}(k,0,0;\zeta,a)$ simplify to

$$\tilde{G}^{|}_{xx}(k;\zeta) = \frac{\kappa_2\kappa_3}{\varepsilon_2\kappa_3+\varepsilon_3\kappa_2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad G^{|a|}_{xx}(k;\zeta,a) = \frac{-\bar{r}_1(1-\bar{r}_2^2)\kappa_2}{2(e^{2a\kappa_3}-\bar{r}_1\bar{r}_2)\varepsilon_2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$\tilde{G}^{|}_{yy}(k;\zeta) = \frac{\zeta^2}{\kappa_2+\kappa_3}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad G^{|a|}_{yy}(k;\zeta,a) = \frac{r_1(1-r_2^2)\zeta^2}{2(e^{2a\kappa_3}-r_1r_2)\kappa_2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$\tilde{G}^{|}_{zz}(k;\zeta) = \frac{-k^2}{\varepsilon_2\kappa_3+\varepsilon_3\kappa_2}\begin{bmatrix} \varepsilon_3/\varepsilon_2 & 1 \\ 1 & \varepsilon_2/\varepsilon_3 \end{bmatrix}, \qquad G^{|a|}_{zz}(k;\zeta,a) = \frac{-\bar{r}_1(1-\bar{r}_2^2)k^2}{2(e^{2a\kappa_3}-\bar{r}_1\bar{r}_2)\kappa_2\varepsilon_3}\begin{bmatrix} \varepsilon_3/\varepsilon_2 & 1 \\ 1 & \varepsilon_2/\varepsilon_3 \end{bmatrix},$$

$$\tilde{G}^{|}_{xz}(k;\zeta) = \frac{ik}{\varepsilon_2\kappa_3+\varepsilon_3\kappa_2}\begin{bmatrix} \varepsilon_3\bar{\kappa}_2 & \kappa_2 \\ -\kappa_3 & -\varepsilon_2\bar{\kappa}_3 \end{bmatrix}, \qquad G^{|a|}_{xz}(k;\zeta,a) = \frac{i\bar{r}_1(1-\bar{r}_2^2)k}{2(e^{2a\kappa_3}-\bar{r}_1\bar{r}_2)}\begin{bmatrix} 1/\varepsilon_2 & 1/\varepsilon_3 \\ 1/\varepsilon_2 & 1/\varepsilon_3 \end{bmatrix},$$

$$\tilde{G}^{|}_{zx}(k;\zeta) = \frac{-ik}{\varepsilon_2\kappa_3+\varepsilon_3\kappa_2}\begin{bmatrix} \varepsilon_3\bar{\kappa}_2 & -\kappa_3 \\ \kappa_2 & -\varepsilon_2\bar{\kappa}_3 \end{bmatrix}, \qquad G^{|a|}_{zx}(k;\zeta,a) = \frac{-i\bar{r}_1(1-\bar{r}_2^2)k}{2(e^{2a\kappa_3}-\bar{r}_1\bar{r}_2)}\begin{bmatrix} 1/\varepsilon_2 & 1/\varepsilon_2 \\ 1/\varepsilon_3 & 1/\varepsilon_3 \end{bmatrix}. \quad \text{(A9)}$$

APPENDIX B: SIGNED CORRELATORS OF THE ROUGHNESS PROFILE

We here obtain the correlation functions of positive and negative components of the roughness profile for a Gaussian generating functional of roughness correlation functions,

$$\langle e^{\int d\mathbf{x}\alpha(\mathbf{x})h(\mathbf{x})}\rangle = e^{\frac{1}{2}\int d\mathbf{x}d\mathbf{y}\alpha(\mathbf{x})D_2(\mathbf{x}-\mathbf{y})\alpha(\mathbf{y})}, \tag{B1}$$

that is fully determined by the two-point correlation function $\langle h(\mathbf{x})h(\mathbf{y})\rangle = D_2(\mathbf{x}-\mathbf{y})$. We in the following assume that $D_2(0) \geq D_2(\mathbf{x}-\mathbf{y}) > 0$.

Exploiting an integral representation of the $x\theta(x)$ distribution, one has that

$$\begin{aligned} h_\pm(\mathbf{x}) &= h(\mathbf{x})\theta(\pm h(\mathbf{x})) \\ &= \pm\frac{1}{2\pi}\lim_{\varepsilon\to 0^+}\int_{-\infty}^{\infty}\frac{d\beta}{(\beta - i\varepsilon)^2}e^{\pm i\beta h(\mathbf{x})} \\ &= \pm\lim_{\varepsilon\to 0^+}\int_0^\infty \lambda d\lambda e^{-\varepsilon\lambda}\int_{-\infty}^{\infty}\frac{d\beta}{2\pi}e^{-i\lambda\beta}e^{\pm i\beta h(\mathbf{x})}. \end{aligned} \tag{B2}$$

We use Eq. (B2) to write

$$\begin{aligned} \langle h_+(\mathbf{x})h_\pm(\mathbf{y})\rangle = \pm\lim_{\varepsilon\to 0^+}\int_0^\infty \lambda_1 d\lambda_1\int_0^\infty \lambda_2 d\lambda_2 e^{-\varepsilon(\lambda_1+\lambda_2)} \\ \times\int\frac{d\beta}{(2\pi)^2}e^{-i\lambda\cdot\beta}\langle e^{i(\beta_1 h(\mathbf{x})\pm\beta_2 h(\mathbf{y}))}\rangle. \end{aligned} \tag{B3}$$

The expectation in Eq. (B3) is of the form given in Eq. (B1) with $\alpha(\mathbf{x}') = i(\beta_1\delta(\mathbf{x}'-\mathbf{x}) \pm \beta_2\delta(\mathbf{x}'-\mathbf{y}))$, and therefore evaluates to

$$\langle e^{i(\beta_1 h(\mathbf{x})\pm\beta_2 h(\mathbf{y}))}\rangle = e^{-\frac{1}{2}\beta^T\cdot M_\pm\cdot\beta}, \tag{B4}$$

where the symmetric, real, and positive 2×2 matrix

$$M_\pm = \begin{bmatrix} D_2(0) & \pm D_2(\mathbf{x}-\mathbf{y}) \\ \pm D_2(\mathbf{x}-\mathbf{y}) & D_2(0)\end{bmatrix} \tag{B5}$$

has the determinant $\det M_\pm = D_2^2(0) - D_2^2(\mathbf{x}-\mathbf{y}) > 0$ for $|\mathbf{x}-\mathbf{y}| > 0$. Performing the two-dimensional Gaussian integral in $\beta = (\beta_1, \beta_2)$ (for $|\mathbf{x}-\mathbf{y}| > 0$) gives

$$\begin{aligned} \langle h_+(\mathbf{x})h_\pm(\mathbf{y})\rangle = \pm\frac{(\det M_\pm)^{-1/2}}{2\pi}\int_0^\infty d\lambda_1 \\ \times\int_0^\infty d\lambda_2\lambda_1\lambda_2 e^{-\frac{1}{2}\lambda^T\cdot M_\pm^{-1}\cdot\lambda}. \end{aligned} \tag{B6}$$

Converting to polar coordinates $(\lambda_1, \lambda_2) = \lambda(\cos\theta, \sin\theta)$ and noting that the integral extends over the first quadrant with $0 < \theta < \pi/2$ only,

$$\begin{aligned} \langle h_+(\mathbf{x})h_\pm(\mathbf{y})\rangle &= \pm\frac{(\det M_\pm)^{-1/2}}{2\pi}\int_0^{\pi/2}d\theta\frac{\sin(2\theta)}{2}\int_0^\infty \lambda^3 d\lambda e^{-\frac{1}{2}\lambda^2(D_2(0)\mp\sin(2\theta)D_2(\mathbf{x}-\mathbf{y}))/\det M_\pm} \\ &= \pm\frac{(\det M_\pm)^{3/2}}{4\pi}\int_0^\pi d\theta\frac{\sin\theta}{(D_2(0)\mp D_2(\mathbf{x}-\mathbf{y})\sin\theta)^2} \\ &= \pm\frac{D_2(0)}{2\pi}\left(\sin\phi + \left(\frac{\pi}{2}\pm\frac{\pi}{2} - \phi\right)\cos\phi\right), \end{aligned} \tag{B7}$$

with $\cos\phi = D_2(\mathbf{x}-\mathbf{y})/D_2(0)$, $0 < \phi < \pi/2$. This result is reproduced in Eq. (31). The last expression uses the fact that the lengths $D_2(0)$, $D_2(\mathbf{x}-\mathbf{y})$, and $\det M_\pm$ can be interpreted as the sides of a right triangle with hypotenuse $D_2(0)$.

APPENDIX C: ANGULAR INTEGRALS

For the class of correlation functions

$$D_s(q) = 2\pi\sigma^2 l_c^2(1 + q^2 l_c^2/(2s))^{-1-s} \quad \text{with} \quad s > 0, \tag{C1}$$

the angular integrals of Eqs. (48), (52), (55a), and (55b) are all of the form

$$\begin{aligned} A_n(s) &= \int_{-\pi}^{\pi}\frac{d\theta\cos^n\theta}{(1+a-b\cos\theta)^{s+1}} \\ &= \frac{\Gamma(s+1-n)}{\Gamma(s+1)}\frac{\partial^n}{\partial b^n}\frac{2\pi}{(1+a+b)^{s+1-n}} \\ &\quad\times {}_2F_1\left(\frac{1}{2}, s+1-n; 1; \frac{2b}{1+a+b}\right), \end{aligned} \tag{C2}$$

with $a = \frac{1}{2}(k^2 + k'^2)l_c^2/s \geq b = kk' l_c^2/s > 0$ and $n = 0, 1, 2$. They are given by values of the generalized hypergeometric function ${}_2F_1(\frac{1}{2}, \nu; 1; x)$ for any $s > 0$.

The exponential roughness correlation D_{Exp} of Eq. (35) corresponds to $s = 1/2$, and the relevant angular integrals in this case are complete elliptic integrals:

$$A_0(1/2) = \frac{4}{(1+a-b)\sqrt{1+a+b}} E\left(\frac{2b}{1+a+b}\right),$$
$$A_1(1/2) = \frac{4}{(1+a-b)b\sqrt{1+a+b}} \left((1+a)E\left(\frac{2b}{1+a+b}\right) - (1+a-b)K\left(\frac{2b}{1+a+b}\right)\right),$$
$$A_2(1/2) = \frac{4}{(1+a-b)b^2\sqrt{1+a+b}} \left((2(1+a)^2-b^2)E\left(\frac{2b}{1+a+b}\right) - 2(1+a)(1+a-b)K\left(\frac{2b}{1+a+b}\right)\right), \quad \text{(C3)}$$

with $a = (k^2+k'^2)l_c^2$ and $b = 2kk'l_c^2$.

The limit $s \to \infty$ of the Gaussian correlation in Eq. (34) is best obtained directly. The angular integrals in this limit are

$$A_n(\infty) = e^{-l_c^2(k^2+k'^2)/2} \int_{-\pi}^{\pi} d\theta \cos^n\theta e^{l_c^2 kk'\cos\theta} = 2\pi e^{-l_c^2(k^2+k'^2)/2} \frac{\partial^n}{\partial\alpha^n} I_0(\alpha)\big|_{\alpha=l_c^2kk'}, \quad \text{(C4)}$$

where $I_0(x)$ is the modified Bessel function of the first kind of zeroth order. The relevant angular integrals for Gaussian roughness correlation thus are

$$A_0(\infty) = 2\pi e^{-l_c^2(k^2+k'^2)/2} I_0(l_c^2kk'),$$
$$A_1(\infty) = 2\pi e^{-l_c^2(k^2+k'^2)/2} I_1(l_c^2kk'),$$
$$A_2(\infty) = \pi e^{-l_c^2(k^2+k'^2)/2} (I_0(l_c^2kk') + I_2(l_c^2kk')). \quad \text{(C5)}$$

APPENDIX D: THE RESPONSE FUNCTION

The roughness correction to the Casimir free energy of order σ^2 is given in Eq. (44). This correction is linear in $D(q)$, and one may define [17] the response function $R_T(q,a)$ of Eq. (79) by

$$\Delta F_T^{\text{Cas}}[a] = \frac{1}{2}\langle \text{Tr}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle - \frac{1}{2}\langle \text{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle + \frac{1}{2}\,\text{Tr}\delta\tilde{\mathbf{V}}^h\mathbf{G}^{|a|} - \frac{1}{4}\langle \tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle = \int_0^\infty \frac{qdq}{2\pi} D(q) R_T(q,a). \quad \text{(D1)}$$

To obtain $R_T(q,a)$, we change the integration variable from $\mathbf{k}'$ to $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ in Eqs. (45), (48), (50), and (52) and choose $\mathbf{k} = (k,0)$ to define the positive x axis. In these coordinates, $k'_x = k + q\cos\theta$, $k'_y = q\sin\theta$, and explicit expressions for the response function $R_T(q,a)$ can be read off from

$$\frac{1}{2}\langle \text{Tr}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle = \int_0^\infty \frac{qdq}{2\pi} D(q) \sum_n (-AT) \int_0^\infty \frac{kdk}{2\pi} \kappa\kappa_\varepsilon \left(\frac{\bar{r}^2}{e^{2a\kappa}-\bar{r}^2} + \frac{r^2}{e^{2a\kappa}-r^2}\right), \quad \text{(D2a)}$$

$$-\frac{1}{2}\langle \text{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle = \int_0^\infty \frac{qdq}{(2\pi)^2} D(q) \sum_n (-AT)(\varepsilon-1)^2 \int_0^\infty \frac{kdk}{2\pi} \int_{-\pi}^{\pi} d\theta \left[\frac{r(1-r^2)\zeta^2}{4(e^{2a\kappa}-r^2)\kappa_\varepsilon} \left(\frac{\kappa'\kappa'_\varepsilon}{\varepsilon\kappa'+\kappa'_\varepsilon}\left(\frac{k'_y}{k'}\right)^2 + \frac{\zeta^2}{\kappa'+\kappa'_\varepsilon}\left(\frac{k'_x}{k'}\right)^2\right) + \frac{\bar{r}(1-\bar{r}^2)}{4(e^{2a\kappa}-\bar{r}^2)\varepsilon} \left(\frac{\varepsilon k^2 k'^2}{\kappa_\varepsilon(\varepsilon\kappa'+\kappa'_\varepsilon)} - kk'_x\bar{r}' - \frac{\kappa_\varepsilon\kappa'\kappa'_\varepsilon}{\varepsilon\kappa'+\kappa'_\varepsilon}\left(\frac{k'_x}{k'}\right)^2 - \frac{\kappa_\varepsilon\zeta^2}{\kappa'+\kappa'_\varepsilon}\left(\frac{k'_y}{k'}\right)^2\right)\right], \quad \text{(D2b)}$$

$$-\frac{1}{4}\langle \text{Tr}\tilde{\mathbf{V}}^h\tilde{\mathbf{G}}^{|a|}\tilde{\mathbf{V}}^h\mathbf{G}^{|a|}\rangle = \int_0^\infty \frac{qdq}{(2\pi)^2} D(q) \sum_n (-AT)(\varepsilon-1)^2 \int_0^\infty \frac{kdk}{2\pi} \int_{-\pi}^{\pi} d\theta \left[\frac{r(1-r^2)\zeta^2}{16(e^{2a\kappa}-r^2)\kappa_\varepsilon} \left(\frac{r'(1-r'^2)\zeta^2}{(e^{2a\kappa'}-r'^2)\kappa'_\varepsilon}\left(\frac{k'_x}{k'}\right)^2 - \frac{2\bar{r}'(1-\bar{r}'^2)\kappa'_\varepsilon}{(e^{2a\kappa'}-\bar{r}'^2)\varepsilon}\left(\frac{k'_y}{k'}\right)^2\right) + \frac{\bar{r}\bar{r}'(1-\bar{r}^2)(1-\bar{r}'^2)}{16(e^{2a\kappa}-\bar{r}^2)(e^{2a\kappa'}-\bar{r}'^2)} \left(\frac{k^2k'^2}{\kappa_\varepsilon\kappa'_\varepsilon} + \frac{2kk'_x}{\varepsilon} + \frac{\kappa_\varepsilon\kappa'_\varepsilon}{\varepsilon^2}\left(\frac{k'_x}{k'}\right)^2\right)\right], \quad \text{(D2c)}$$

$$\frac{1}{2}\,\text{Tr}\delta\tilde{\mathbf{V}}\mathbf{G}^{|a|} = \int_0^\infty \frac{qdq}{2\pi} D(q) \sum_n AT(\varepsilon-1)^2 \int_0^\infty \frac{kdk}{2\pi} \left[\frac{\bar{r}(1-\bar{r}^2)}{4(e^{2a\kappa}-\bar{r}^2)} \frac{k^2q^2}{\kappa_\varepsilon(\varepsilon\kappa'+\kappa'_\varepsilon)} + \left(\frac{r(1-r^2)\zeta^2}{4(e^{2a\kappa}-r^2)\kappa_\varepsilon} - \frac{\bar{r}(1-\bar{r}^2)\kappa_\varepsilon}{4(e^{2a\kappa}-\bar{r}^2)\varepsilon}\right)\left(\frac{\kappa'\kappa'_\varepsilon/2}{\varepsilon\kappa'+\kappa'_\varepsilon} + \frac{\zeta^2/2}{\kappa'+\kappa'_\varepsilon} - \frac{g^2\zeta}{1+\sqrt{\varepsilon}}\right)\right]. \quad \text{(D2d)}$$

In the last (counterterm) expression of Eq. (D2d), $\kappa' = \sqrt{\mathbf{q}^2+\zeta^2}$ and $\kappa'_\varepsilon = \sqrt{\mathbf{q}^2+\zeta^2\varepsilon(\zeta)}$. Note that the angular integration in these coordinates cannot be performed analytically.

[1] H. B. G. Casimir, Kon. Ned. Akad. Wetensch. Proc. **51**, 793 (1948).
[2] E. M. Lifshitz, Zh. Eksp. Teor. Fiz. **29**, 94 (1955); Sov. Phys. JETP **2**, 73 (1956).
[3] J. Schwinger, L. L. DeRaad, and K. Milton, Ann. Phys. (N.Y.) **115**, 1 (1978).
[4] Lowell S. Brown and G. Jordan Maclay, Phys. Rev. **184**, 1272 (1969).
[5] B. V. Derjaguin and I. I. Abrikossova, J. Phys. Chem. Solids **5**, 1 (1958).
[6] S. K. Lamoreaux, Phys. Rev. Lett. **78**, 5 (1997).
[7] A. A. Maradudin and P. Mazur, Phys. Rev. B **22**, 1677 (1980); P. Mazur and A. A. Maradudin, Phys. Rev. B **23**, 695 (1981).
[8] M. Yu. Novikov, A. S. Sorin, and V. Ya. Chernyak, Theor. Math. Phys. **82**, 124 (1990); **82**, 252 (1990); **91**, 658 (1992); **92**, 773 (1992).
[9] U. Mohideen and A. Roy, Phys. Rev. Lett. **81**, 4549 (1998).
[10] R. S. Decca, D. López, E. Fischbach, and D. E. Krause, Phys. Rev. Lett. **91**, 050402 (2003); D. E. Krause, R. S. Decca, D. López, and E. Fischbach, Phys. Rev. Lett. **98**, 050403 (2007).
[11] P. J. van Zwol, G. Palasantzas, and J. Th. M. De Hosson, Appl. Phys. Lett. **91**, 144108 (2007); P. J. van Zwol, G. Palasantzas, M. van de Schootbrugge, and J. Th. M. De Hosson, Appl. Phys. Lett. **92**, 054101 (2008).
[12] P. J. van Zwol, G. Palasantzas, and J. Th. M. De Hosson, Phys. Rev. B **77**, 075412 (2008).
[13] W. Broer, G. Palasantzas, and J. Knoester, Phys. Rev. B **85**, 155410 (2012).
[14] G. L. Klimchitskaya, A. Roy, U. Mohideen, and V. M. Mostepanenko, Phys. Rev. A **60**, 3487 (1999).
[15] T. Emig, A. Hanke, R. Golestanian, and M. Kardar, Phys. Rev. Lett. **87**, 260402 (2001).
[16] R. Büscher and T. Emig, Phys. Rev. A **69**, 062101 (2004).
[17] P. A. M. Neto, A. Lambrecht, and S. Reynaud, Phys. Rev. A **72**, 012115 (2005).
[18] C. Genet, A. Lambrecht, P. Maia Neto, and S. Reynaud, Europhys. Lett. **62**, 484 (2003); P. A. Maia Neto, A. Lambrecht, and S. Reynaud, Europhys. Lett. **69**, 924 (2005); P. A. Maia Neto, A. Lambrecht, and S. Reynaud, J. Phys. A **39**, 6517 (2006); A. Lambrecht, P. A. Maia Neto, and S. Reynaud, New J. Phys. **8**, 243 (2006).
[19] G. Palasantzas and J. Th. M. De Hosson, Phys. Rev. B **72**, 115426 (2005).
[20] M. Bordag, G. L. Klimchitskaya, U. Mohideen, and V. M. Mostepanenko, *Advances in the Casimir Effect* (Oxford University Press, New York, 2009).
[21] C. D. Fosco, F. C. Lombardo, and F. D. Mazzitelli, Phys. Rev. D **84**, 105031 (2011).
[22] H. Y. Wu and M. Schaden, Phys. Rev. D **85**, 045008 (2012).
[23] C. D. Fosco, F. C. Lombardo, and F. D. Mazzitelli, Phys. Rev. D **86**, 125018 (2012).
[24] H. B. Chan, Y. Bao, and J. Zou, Phys. Rev. Lett. **101**, 030401 (2008).
[25] F. Intravaia, S. Koev, W. Jung, A. A. Talin, P. S. Davids, R. S. Decca, V. A. Aksyuk, D. A. R. Dalvit, and D. Lopez, Nat. Commun. **4**, 2515 (2013).
[26] J. Gomis and S. Weinberg, Nucl. Phys. **B469**, 473 (1996).
[27] Steven Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, Cambridge, United Kingdom, 2005).
[28] H. Leutwyler, Ann. Phys. (N.Y.) **235**, 165 (1994).
[29] D. Deutsch and P. Candelas, Phys. Rev. D **20**, 3063 (1979).
[30] N. Graham, R. L. Jaffe, V. Khemani, M. Quandt, M. Scandurra, and H. Weigel, Nucl. Phys. **B645**, 49 (2002); E. Elizalde, J. Phys. A **36**, L567 (2003); N. Graham, R. L. Jaffe, V. Khemani, M. Quandt, O. Schrder, and H. Weigel, Nucl. Phys. **B677**, 379 (2004); K. A. Milton, J. Wagner, and K. Kirsten, Phys. Rev. D **80**, 125028 (2009); S. A. Fulling, Int. J. Mod. Phys. A **25**, 2364 (2010); N. Graham, A. Shpunt, T. Emig, S. J. Rahi, R. L. Jaffe, and M. Kardar, Phys. Rev. D **83**, 125007 (2011).
[31] H. M. Fried, *Functional Methods and Models in Quantum Field Theory* (MIT Press, Cambridge, MA, 1972); Peter Becher, Manfred Böhm, and Hans Joos, *Gauge Theories of Strong and Electroweak Interactions* (Wiley, New York, 1984); J. I. Kapusta and C. Gale, *Finite-Temperature Field Theory* (Cambridge University Press, Cambridge, United Kingdom, 2006).
[32] T. Emig, A. Hanke, R. Golestanian, and M. Kardar, Phys. Rev. A **67**, 022114 (2003); I. Cavero-Pelaez, K. A. Milton, P. Parashar, and K. V. Shajesh, Phys. Rev. D **78**, 065018 (2008); **78**, 065019 (2008).
[33] F. S. S. Rosa, D. A. R. Dalvit, and P. W. Milonni, Phys. Rev. A **78**, 032117 (2008).
[34] A. Gusso and U. B. Reis, Europhys. Lett. **99**, 36003 (2012).
[35] K. V. Shajesh and M. Schaden, Phys. Rev. D **83**, 125032 (2011).
[36] J. Daillant and A. Gibaud, *X-Ray and Neutron Reflectivity: Principles and Applications* (Springer, New York, 1999).
[37] Heinz Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings* (Springer-Verlag, Berlin, 1988).
[38] A. Lambrecht and V. N. Marachevsky, Phys. Rev. Lett. **101**, 160403 (2008).
[39] F. Intravaia, P. S. Davids, R. S. Decca, V. A. Aksyuk, D. A. R. Dalvit, and D. Lopez, Phys. Rev. A **86**, 042101 (2012).