

**Gauge-invariant theory of truncated quantum light-matter interactions in arbitrary media**Chris Gustin <sup>1,\*</sup>, Sebastian Franke <sup>2,3</sup> and Stephen Hughes <sup>2</sup><sup>1</sup>*Edward L. Ginzton Laboratory, Stanford University, Stanford, California 94305, USA*<sup>2</sup>*Department of Physics, Engineering Physics, and Astronomy, Queen's University, Kingston, Ontario, Canada K7L 3N6*<sup>3</sup>*Technische Universität Berlin, Institut für Theoretische Physik, Nichtlineare Optik und Quantenelektronik, Hardenbergstraße 36, 10623 Berlin, Germany*

(Received 24 August 2022; accepted 23 December 2022; published 27 January 2023)

The loss of gauge invariance in models of light-matter interaction which arises from material and photonic space truncation can pose significant challenges to conventional quantum optical models when matter and light strongly hybridize. In structured photonic environments, necessary in practice to achieve strong light-matter coupling, a rigorous model of field quantization within the medium is also needed. Here we use the framework of macroscopic QED by quantizing the fields in an arbitrary material system, with a spatially dependent dispersive and absorptive dielectric, starting from a fundamental light-matter action. We truncate the material and mode degrees of freedom while respecting the gauge principle by imposing a partial gauge-fixing constraint during canonical quantization, which admits a large number of gauges including the Coulomb and multipolar gauges commonly used in quantum optics. We also consider gauge conditions with explicit time dependence, enabling us to unambiguously introduce additional phenomenologically time-dependent light-matter interactions in any gauge. Our results allow one to derive rigorous nonrelativistic models of ultrastrong light-matter interactions in structured photonic environments with no gauge ambiguity. Results for two-level systems and the dipole approximation are discussed, as well as how to go beyond the dipole approximation for effective single-particle models. By comparing with the limiting case of an inhomogeneous dielectric, where dispersion and absorption can be neglected and the fields can be expanded in terms of the generalized transverse eigenfunctions of the dielectric, we show how lossy systems can introduce an additional gauge ambiguity, which we resolve and predict to have fundamental implications for open quantum system models. Finally, we show how observables in mode-truncated systems can be calculated without ambiguity by using a simple gauge-invariant model of photodetection.

DOI: [10.1103/PhysRevA.107.013722](https://doi.org/10.1103/PhysRevA.107.013722)**I. INTRODUCTION**

In nanophotonics, one often would like to describe the interaction of a small number of emitters, treated as microscopic degrees of freedom, interfacing via the electromagnetic field with a macroscopic medium, wherein the different degrees of freedom are not tracked explicitly. In classical electromagnetism, this is accomplished by the macroscopic Maxwell's equations, where the medium is, assuming a linear response of the medium to applied fields, ascribed a dielectric function, which in general can be frequency-dependent (allowing for dispersion) and complex (allowing for energy losses via absorption). In fact, the requirement of the constituent medium response to the applied field to follow a causal relationship implies, via the Kramers-Kronig relations, that in general a frequency-dependent dielectric function is complex, and vice versa.

In quantum mechanics, the direct quantization of the macroscopic Maxwell's equations is complicated by the fact that under an imaginary permittivity, the operators describing the electromagnetic field would in general decay to zero

amplitude, in violation of the fundamental commutation relations [1]. One particularly successful method to quantize the electromagnetic field in a dispersive and absorbing medium is the macroscopic quantum electrodynamics (QED) approach [1–8], where the field is expanded in terms of the photonic Green's function of the medium (obtained from the impulse response of the electric field to a localized dipole source) and a bosonic polariton field. As the Green's function can be obtained by purely classical calculations—including analytic solutions for simple geometries and general numerical techniques (e.g., finite-difference time-domain simulations) for general cases—this powerful approach provides a quantum mechanical framework for studying the dynamics of light-matter systems in practical nanophotonic settings. The Green's function quantization can be regarded as a generalization of the usual normal mode expansion from lossless systems. Moreover, the Green's functions can also be obtained through mode expansion techniques, even for complex geometries and in the presence of photon loss [9,10].

The macroscopic QED formalism can be justified on purely phenomenological grounds, by virtue of its simultaneous fulfillment of the macroscopic Maxwell's equations and Lorentz equations of motion, the fundamental quantization commutation relations of the electromagnetic field, and the

\*cgustin@stanford.edu

dissipation-fluctuation theorem [4]. Microscopic derivations can also be performed, wherein a material medium “reservoir” field is coupled via the fundamental minimal coupling prescription of QED to the vacuum electromagnetic fields, and a process of “Fano diagonalization” is used to diagonalize the total Hamiltonian (medium plus electromagnetic field) in terms of bosonic polariton operators which can then couple to microscopic material particles [3,11]. In this latter microscopic derivation, the coupling can be written as a quantized Hamiltonian interaction, or more generally quantization can be performed at the level of an initial Lagrangian [7] by identifying canonical variables and quantizing in accordance with Dirac’s prescription for quantization with constraints [12]. These results are also in accordance with the phenomenological quantization approach.

Macroscopic QED, already having proved a powerful theoretical tool for modeling light-matter interactions in a medium (e.g., spontaneous emission in arbitrary environments [13–15], dipole-dipole interactions [16,17], Casimir-Polder forces [18], and surface-enhanced Raman spectroscopy [19]), is an excellent candidate for improving models and providing fundamental analysis of light-matter interactions in the so-called ultrastrong coupling (USC) regime [20,21]. In the USC regime, the parameters which characterize the coupling between the material and field degrees of freedom become substantial compared to the bare resonances of the subsystems, which strongly hybridizes the field and matter degrees of freedom, and common frameworks for understanding the dynamics of the interaction break down. These changes can be dynamical, as a consequence of having to forgo the widely used rotating-wave approximation, but recently it has come to be more widely understood that the fundamental Hamiltonian used to describe the coupling between the subsystems itself can become questionable in any situation where the field or material degrees of freedom are expressed in a *truncated basis*. In particular, the gauge invariance of the theory is broken when the minimal coupling Hamiltonian is expressed in terms of variables which are truncated to a finite energy basis for the material degrees of freedom, or the number of modes (and presumably Fock number states) for the field [22–24]. Notably, the loss of gauge invariance remains highly significant in interaction regimes where the truncation process retains all energy levels near-resonant with the ultrastrong interaction, where a truncated model should, in principle, be accurate and independent of choice of gauge.

The breakdown of gauge-invariance in material systems [23,25–29] can be understood by noting that a truncation in an energy basis [e.g., to the widely used two-level system (TLS) model] implies a truncation in a position basis (which is continuous without truncation). As a U(1) gauge theory, QED promotes the global symmetry of the Schrödinger equation’s invariance under a total change in the phase of the state vector to a *local* symmetry. The presence of this new symmetry induces a minimal coupling to a gauge boson field. In the case of nonrelativistic QED, a transformation of a wave function of a particle with charge  $q$ , through

$$\psi(\mathbf{x}) \rightarrow \exp[iq\Lambda(\mathbf{x})/\hbar]\psi(\mathbf{x}) \quad (1)$$

can be compensated by a corresponding gauge transformation of the potential fields,  $\mathbf{A} \rightarrow \mathbf{A} + \nabla\Lambda$  and  $\phi \rightarrow \phi - \dot{\Lambda}$ . All

physical results must be invariant under local U(1) transformations.

Critically, if the model is to be implemented in a truncated basis, the truncation must be carried out in a way which is consistent with the gauge transformation in Eq. (1); that is, a gauge transformation of the field must be able to be compensated with an appropriate dimensional unitary transformation of the state vector. In a different context, these insights form the basis of lattice gauge theory, introduced by Wilson [30] to study quantum chromodynamics on a lattice, where the continuous representation of position is truncated to a finite basis in a manner which respects the gauge symmetry of the theory. Recent work has shown that gauge invariance can be restored in truncated material systems [27,29,31], by using a generalized minimal coupling replacement in the form of a unitary transformation, which correctly constrains the light-matter interaction within the truncated subspace.

Equivalent insights can be used to show that truncation of photonic degrees of freedom (e.g., the number of photonic modes) also yields gauge-dependent predictions [22,24], and that gauge invariance can be restored by a similar unitary transformation. It has also been noted in different contexts, such as high-order harmonic generation [32] and tight-binding models [33], that truncation in the Coulomb gauge (and its analogs) allows for converging results with less modes, while the multipolar gauge (and its analogs) allows for less material states. In the context of media with loss, discrete mode expansions of the electromagnetic fields can take the form of, e.g., quasimodes of various types [34–36] or quasinormal modes (QNMs) [10,37,38]—where for the latter, a fully quantized theory has been developed recently [39] and applied to plasmonic single-photon sources [40] and coupled resonators [41], including gain-loss systems [42,43], and appears to be an excellent candidate for modeling ultrastrong cavity-QED interactions in realistic photonic media.

In this work, we take the view that any rigorous picture of open system quantum optics that involves a few discrete modes [44] manifests in some degree of mode truncation, and thus understanding gauge invariance in these models is essential.

In the quantization of systems with matter (that is, degrees of freedom beyond the “passive” material medium considered in the macroscopic QED approach) an additional gauge symmetry beyond that of the field potentials arises due to the polarization field of the matter component. In this manner, for light-matter interactions it is therefore necessary to formulate a theoretical framework which allows for gauge transformations consistent with this greater class of symmetry under material truncation. Previous work has shown these methods to maintain gauge invariance on the basis of semiclassical arguments [25,27,31], which is generally sufficient, as after quantization, the gauge is invariably at least partially fixed by the requirement of constraints to quantize the electromagnetic field (and the same approach can be justified without reference to gauge transformations by instead appealing to the notion of constraining interactions within the correctly truncated subspace [29]). However, this approach does not shed insight into what exactly the realizable gauges are and what a gauge transformation consists of in the truncated space. The usual approach is to quantize in the Coulomb gauge and construct

the multipolar gauge by means of the unitary Power-Zienau-Woolley (PZW) transformation [45,46]; however, because a unitary transformation from a fixed gauge cannot implement a *generic* gauge transformation [47], this has raised questions [48] (and the resolutions to these questions [49–51]) about the validity of such a procedure recently. To be consistent with macroscopic QED, any approach must start from a Lagrangian which respects this gauge symmetry of both the electromagnetic and material polarization fields, and contains the material reservoir fields which describe the medium.

In this paper we accomplish this task by quantizing the electromagnetic fields in the presence of both a material medium reservoir field (the “passive” component) and free charged particles (the “active” component), the latter of which can interact with the electromagnetic fields with arbitrary strength, allowing one to study USC effects. Using a *c*-number quantization function method by Woolley [47], we *quantize in a way which does not require choosing a specific gauge, allowing us to derive manifestly gauge-invariant models* which incorporate a broad class of gauges, including the most commonly used gauges in quantum optics: the Coulomb and multipolar (dipole, when using a dipole approximation) gauges. We show explicitly the validity of previous theoretical works on restoring gauge invariance, and shed light on the nonrelativistic limit of their application.

Our results provide a rigorous and gauge-invariant framework for describing light-matter interactions in the USC regime from a first-principles approach, for arbitrary media, and one that can still take advantage of a reduced description of the medium in terms of a linear susceptibility function. We stress that from a theoretical perspective, our results need not be implemented in the context of a macroscopically quantized medium, and indeed the formalism also applies for free space quantization, but the presence of a medium (generally one that supports resonant modes) is necessary to reach the USC regime in practice in optical systems.

Our work also lays the necessary groundwork for the future development of first-principles models of loss from cavity-QED systems in the USC regime. This is timely and highly desired, as it has been shown recently that the nearly universally used phenomenological model of dissipation (standard input-output theory [52]) is insufficient in the USC regime [53]. We expect our work to be useful and applicable to, in addition to quantized QNMs, studies based on, for example, pseudomodes [54,55] or simulations involving matrix product states [56,57].

In addition to laying out the fundamental theory of gauge-invariant interactions in quantum light-matter systems in a quantized and arbitrary medium, our work also contributes three additional main findings:

(i) Any open quantum systems approach to photon loss (e.g., a master equation) in a system interacting ultrastrongly with matter, from a rigorous theoretical perspective, should be derived in the Coulomb gauge, as it is the unique gauge in which the reservoir can be described by a subspace *unentangled* with the light-matter system. However, a reduced notion of a gauge transformation can be defined only with respect to a truncated field (e.g., a cavity mode in cavity-QED models), which has allowed for previous developments of gauge-invariant models [53,58] (in these cases, assuming

phenomenological models of system-reservoir coupling). This necessarily requires a mode-truncated description of the reduced gauge transformation, and thus we propose that there exists a potential intrinsic gauge ambiguity due to mode truncation in rigorous open quantum system models of photon loss, for which the techniques described in this work to retain gauge invariance are important.

(ii) Contrasting previous claims [60], we show that it is possible to introduce unambiguous phenomenological time-dependent light-matter interactions in *any gauge*, provided the time dependence of the gauge condition is consistently accounted for in the quantization.

(iii) By considering an explicit model of photodetection in a truncated mode system, we resolve a gauge ambiguity regarding observables and provide justification for the recent approach that has been used in previous works, for simple model systems of cavity-QED [53,58,59,61]. By identifying the *correctly mode-truncated electric field operator*, we also refute recent claims [62] that the modal expansion operators of the transverse electric field are not the correct operators to couple to external reservoir modes in the case of open quantum systems.

The rest of our paper is organized as follows. In Sec. II we present the fundamental *action* from which we derive Maxwell’s equations and the Lorentz force law in a dispersive and absorbing dielectric, where we treat the medium degrees of freedom explicitly as a frequency and spatially dependent reservoir with an oscillator field.

In Sec. III we show how this general system can be quantized using Dirac’s method of canonical quantization with constraints, using Woolley’s [47] quantization function approach. Following previous works [7,11], we then perform a Fano diagonalization to express part of the quantum Hamiltonian for this (arbitrary-gauge) system as a bosonic polariton harmonic oscillator field, which removes any explicit reference to the medium oscillator degrees of freedom, and express the electromagnetic fields in terms of the photonic Green’s function of the medium and the polariton operators.

In Sec. IV we show how gauge invariance manifests in the quantum theory with the quantization function approach, and show how material and mode truncation can be introduced in a manner that respects the gauge principle. We contrast the form of discrete mode expansion that can be done in a lossy system with the more commonly employed normal mode expansion (namely, using lossless eigenmodes), which is generally only strictly possible in lossless and dispersionless media. Figure 1 shows a schematic conceptual representation of our approach.

In Sec. V we generalize the quantization function approach to quantizing in an arbitrary gauge by allowing the gauge condition to have explicit time dependence. This permits for a broader set of gauge transformations to be considered than have previously appeared in the literature. We then use this formalism to show how phenomenological time-dependent interactions in ultrastrong transverse light-matter interactions can be introduced *unambiguously* in any gauge.

In Sec. VI we apply our results to some common approximations in quantum optics, specifically the dipole and material TLS approximations, and show how to go beyond the dipole approximation for the case of an effective single-particle model.

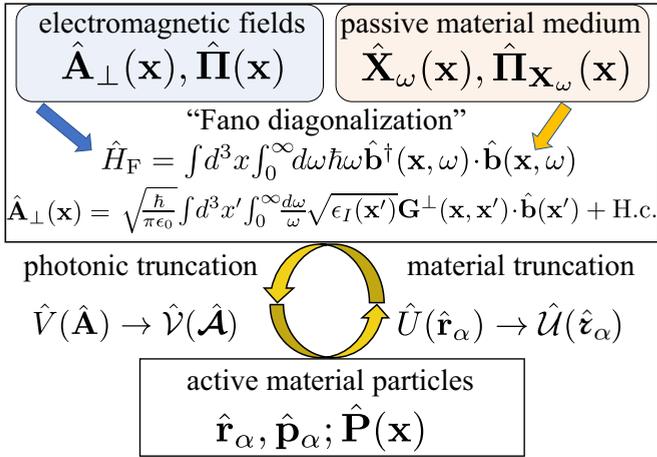


FIG. 1. Schematic representation of the gauge-invariant approach to light-matter interactions within a linear medium (see text for definition of mathematical symbols). The electromagnetic fields interact weakly with a passive material reservoir representing the dispersive and absorbing dielectric medium, allowing for the fields to be expressed as a linear functional of fundamental polariton operators  $\hat{\mathbf{b}}(\mathbf{x}, \omega)$ ,  $\hat{\mathbf{b}}^{\dagger}(\mathbf{x}, \omega)$  (we have suppressed the  $\omega$  functional dependence in the dielectric function, Green’s function, and annihilation operator in the schematic for brevity). Light-matter interactions are introduced by a minimal coupling scheme of free “active” material particles with the (vacuum) electromagnetic fields. Gauge invariance under active material truncation is retained by expressing the Hamiltonian in terms of the unitary operators  $\hat{U}(\hat{\mathbf{e}}_{\alpha})$ , where  $\hat{\mathbf{e}}_{\alpha}$  are the directly truncated position operators of the active material component. Similarly, gauge invariance is retained under photonic truncation by expressing the Hamiltonian in terms of  $\hat{V}(\hat{\mathcal{A}})$ , where  $\hat{\mathcal{A}}$  is the truncated vector potential. The approach is manifestly gauge invariant as the longitudinal component of the vector potential  $\hat{\mathcal{A}}_{\parallel}$  and transverse component of the polarization  $\hat{\mathcal{P}}_{\perp}$  are determined by the arbitrary transverse quantization function  $\mathbf{K}_{\perp}(\mathbf{x}, \mathbf{x}')$ .

In Sec. VII we identify and resolve a potential gauge ambiguity regarding observables of the electromagnetic field, by introducing an explicit simple model of photodetection in a truncated mode system. Using our formalism, we show that for mode truncation to be correctly performed, it must be done with respect to the *vector potential*, and that the correctly mode-truncated form of the electric field operator subsequently takes a modified form.

Finally, in Sec. VIII we conclude. In addition, we include four Appendixes. Appendixes A and B give extra details on the fundamental light-matter interaction action and canonical quantization procedure in the presence of constraints. In Appendix C we show in more detail how a discrete mode expansion can be generally constructed from the continuous polariton operators and connect to the important case of quantized QNMs. We consider in Appendix D the case of an inhomogeneous but nondispersive and lossless dielectric, with a real dielectric permittivity that is independent of frequency  $\epsilon(\mathbf{x})$ . In this important special case (justifiable in a limited frequency regime), the field variables can be expressed in terms of the so-called generalized transverse eigenfunctions of the system, which are normal modes of Maxwell’s equations in the dielectric medium with closed or periodic

boundary conditions, by choosing the generalized Coulomb gauge, which satisfies  $\nabla \cdot [\epsilon(\mathbf{x})\mathbf{A}(\mathbf{x})] = 0$ , or the generalized multipolar gauge. We quantize the electromagnetic field in this medium using Dirac’s constrained quantization technique for the generalized Coulomb gauge condition, which allows us to recover previously known results (e.g., Refs. [63–65]), now using a systematic method. We then discuss how gauge-invariant truncated models can be obtained, similarly to the procedure in the main text.

## II. ELECTROMAGNETIC FIELDS, ARBITRARY DIELECTRIC PASSIVE MEDIUM, AND ACTIVE MATERIAL PARTICLES

We take our model to consist of the electromagnetic field, a passive medium reservoir field  $\mathbf{X}_{\omega}(\mathbf{x}, t)$  (corresponding to “bound” charges), as well as active free particles, indexed by  $\alpha$ , with charge  $q_{\alpha}$ , mass  $m_{\alpha}$ , and position coordinate  $\mathbf{r}_{\alpha}(t)$ . For simplicity we assume no net charge, such that  $\sum_{\alpha} q_{\alpha} = 0$ . We would like to ultimately model the interaction of the free charges with the electromagnetic field using a *macroscopic* approach, where the field can be expressed in a quantized form, using the photonic Green’s function of the medium, which is determined by its dielectric function  $\epsilon(\mathbf{x}, \omega)$ . In general, the dielectric constant is a complex-valued function that depends on space and frequency with real and imaginary parts  $\epsilon_R(\mathbf{x}, \omega)$  and  $\epsilon_I(\mathbf{x}, \omega)$ , respectively. It may also be anisotropic, but we will consider this function as a scalar.

We consider the action  $\mathcal{S} = \mathcal{S}[A_{\mu}; \mathbf{r}_{\alpha}; \mathbf{X}_{\omega}]$ . This action is a functional of the electromagnetic potential fields  $A^{\mu}(x) = (\phi(\mathbf{x}, t)/c, \mathbf{A}(\mathbf{x}, t))$ , the medium field excitations  $\mathbf{X}_{\omega}(\mathbf{x}, t)$ , and the material particle coordinates  $\mathbf{r}_{\alpha}(t)$ , as well as the time derivatives of each of these respective quantities. The medium is described by a dielectric function with real and imaginary parts  $\epsilon_R(\mathbf{x}, \omega)$  and  $\epsilon_I(\mathbf{x}, \omega)$ , describing dispersion and absorption of the medium, respectively, and which satisfy the Kramers-Kronig relations. We assume nonmagnetic media, such that  $\mu(\mathbf{x}, \omega) \approx 1$ , although the theory can be generalized to also incorporate a general complex magnetic susceptibility [7]. Note that here and throughout,  $\omega$  refers to a continuous modal index, and not the argument of a Fourier transform of time. The full action is given in Appendix A, as well as the equations of motion it generates. These are the Ampère and Lorentz Eqs. (A5b) and (A5c), as well as an equation of motion for the passive medium reservoir oscillator field (A5d) and Gauss’s law (A5a), which acts as a constraint.

## III. QUANTIZATION

In order to perform canonical quantization and promote field coordinates to operators on a Hilbert space, we first move to a Hamiltonian picture. To do so, we choose the Lagrangian  $\mathcal{L}$ , which satisfies  $\mathcal{S} = \int dt \mathcal{L}$ , to take the form of the time integrands of the action in Eqs. (A2a)–(A2d). A total time derivative can be added to the Lagrangian without changing the resulting equations of motion, which is related to gauge symmetry of the theory, discussed in Sec. III A. We

then identify the conjugate momenta to the field and particle coordinates:

$$\Pi_\phi = \frac{\delta \mathcal{L}}{\delta \dot{\phi}} = 0, \quad (2a)$$

$$\Pi_{\mathbf{A}} = \frac{\delta \mathcal{L}}{\delta \dot{\mathbf{A}}} = -\epsilon_0 \mathbf{E} - \int_0^\infty d\omega \alpha(\mathbf{x}, \omega) \mathbf{X}_\omega, \quad (2b)$$

$$\mathbf{p}_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_\alpha} = m_\alpha \dot{\mathbf{r}}_\alpha + q_\alpha \mathbf{A}(\mathbf{r}_\alpha), \quad (2c)$$

$$\Pi_{\mathbf{X}_\omega} = \frac{\delta \mathcal{L}}{\delta \dot{\mathbf{X}}_\omega} = \dot{\mathbf{X}}_\omega, \quad (2d)$$

and  $\alpha(\mathbf{x}, \omega)$  is defined in Eq. (A3).

The vanishing of  $\Pi_\phi$  is indicative of the fact that the equations of motion have redundant degrees of freedom; while the four-potential contains four degrees of freedom, only three dynamical equations of motion are given by Eq. (A5b). In the process of canonical quantization, the usual approach of promoting Poisson brackets to commutators of operators requires a one-to-one correspondence for dynamical equations of motion and unconstrained degrees of freedom. Instead, here we have a constrained system, where the longitudinal component of  $\Pi_{\mathbf{A}}$  is constrained by Gauss's law [Eq. (A5a)], and the canonical momentum  $\Pi_\phi$  is constrained to vanish. These constraints restrict the phase space manifold in which the system is to be quantized, and describe a so-called ‘‘singular’’ system, where the Hessian of the Lagrangian does not have full rank [12,66,67].

In addition to the two constraints described above, which we can write as

$$\chi_0 = \Pi_\phi = 0, \quad (3a)$$

$$\chi_1 = \nabla \cdot \Pi_{\mathbf{A}} + \rho_A = 0, \quad (3b)$$

where  $\rho_A$  is the active particle charge density given by Eq. (A4a), a third constraint on the four-potential is required to ensure that one dynamical equation of motion exists for each of the unconstrained quantization variables. The reason for this is that typically in constrained systems, one can quantize by using Dirac's prescription, which modifies the Poisson brackets as to include the effect of the constraints in a manner that implies there exist unconstrained variables which can be quantized in accordance with the usual Poisson bracket prescription [12]. The difficulty here is that we find that the Poisson brackets of the above constraints vanish, which precludes one from directly applying this procedure. This is ultimately because Eq. (3b) cannot be solved to uniquely eliminate one of the field degrees of freedom (e.g.,  $\phi$ ) by expressing it in terms of the others, which is a consequence of the gauge freedom of the theory [that is, Eq. (3b) allows one only to solve for the longitudinal part of  $\Pi_{\mathbf{A}}$ ] [12]. To remedy this, we will apply another constraint, which we will choose to manifestly preserve the gauge symmetry of the Lagrangian, allowing us to quantize in a (mostly) *arbitrary gauge*.

#### A. Gauge symmetry

Note that under a general gauge transformation:

$$\mathbf{A}(\mathbf{x}, t) \rightarrow \mathbf{A}(\mathbf{x}, t) + \nabla \Lambda(\mathbf{x}, t), \quad (4a)$$

$$\phi(\mathbf{x}, t) \rightarrow \phi(\mathbf{x}, t) - \dot{\Lambda}(\mathbf{x}, t), \quad (4b)$$

the Lagrangian is not invariant, while the Eqs. of motion (A5) are. However, if we add the additional term

$$\mathcal{L} \rightarrow \mathcal{L} - \frac{d}{dt} \int d^3x \mathbf{A}(\mathbf{x}, t) \cdot \mathbf{P}(\mathbf{x}, t), \quad (5)$$

where the auxiliary polarization field  $\mathbf{P}$  satisfies  $\nabla \cdot \mathbf{P}(\mathbf{x}, t) = -\rho_A(\mathbf{x}, t)$  (which fixes its longitudinal part, but leaves the transverse part arbitrary), then the Lagrangian becomes invariant. Note here we have used the relation  $\nabla \cdot \mathbf{J}_A + \dot{\rho}_A = 0$ , where  $\mathbf{J}_A$  is the active particle current density given by Eq. (A4b). Also note that the action involves only the vector potential in the free particle part, whereas the medium is coupled to the (manifestly gauge-invariant) electric field. More fundamentally, this is a consequence of the fact the medium part of the action is chosen to give the macroscopic Maxwell's equations, in the form of a dielectric function.

Since we know how to quantize in the Coulomb gauge (where the vector potential is transverse), it would be useful to have a representation of the vector potential in terms of its transverse part  $\mathbf{A}_\perp$ , which is gauge-invariant. To do this, we use a method devised by Woolley, in which the longitudinal component of the vector potential is determined by a  $c$ -number function projected onto the transverse vector potential [47,51]. We call this method the *quantization function* approach.

Note that we can write, following Helmholtz's theorem, the vector potential as  $\mathbf{A} = \mathbf{A}_\perp + \mathbf{A}_\parallel$ , where

$$\mathbf{A}_\parallel(\mathbf{x}, t) = \nabla^x \int d^3x' \mathbf{K}_\parallel(\mathbf{x}, \mathbf{x}') \cdot \mathbf{A}(\mathbf{x}', t), \quad (6)$$

and  $\mathbf{K}_\parallel(\mathbf{x}', \mathbf{x})$  is the Green's function for the divergence operator

$$\mathbf{K}_\parallel(\mathbf{x}, \mathbf{x}') = -\nabla^x \frac{1}{4\pi|\mathbf{x} - \mathbf{x}'|}, \quad (7)$$

which satisfies  $\nabla^x \cdot \mathbf{K}_\parallel(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}')$ , and we have let the  $\parallel$  ( $\perp$ ) subscript denote the longitudinal (transverse) part of a function with respect to its first spatial argument. Equation (6) is verified easily by noting that the gradient of  $\mathbf{K}_\parallel$  gives the longitudinal Dirac delta function

$$\nabla^x \mathbf{K}_\parallel(\mathbf{x}, \mathbf{x}') = \delta^\parallel(\mathbf{x} - \mathbf{x}') = \int \frac{d^3k}{(2\pi)^3} \hat{\mathbf{k}} \hat{\mathbf{k}} e^{ik \cdot (\mathbf{x} - \mathbf{x}')}. \quad (8)$$

Next, by imposing the constraint

$$\chi_2 = \int d^3x' \mathbf{A}(\mathbf{x}', t) \cdot \mathbf{K}(\mathbf{x}', \mathbf{x}) = 0, \quad (9)$$

and noting that  $\mathbf{K}_\parallel(\mathbf{x}, \mathbf{x}')$  is antisymmetric with respect to exchange of its arguments, it follows that one can write

$$\mathbf{A}(\mathbf{x}, t) = \mathbf{A}_\perp(\mathbf{x}, t) + \nabla^x \int d^3x' \mathbf{A}_\perp(\mathbf{x}', t) \cdot \mathbf{K}_\perp(\mathbf{x}', \mathbf{x}), \quad (10)$$

such that the longitudinal part of the vector potential becomes  $\mathbf{A}_\parallel(\mathbf{x}, t) = \nabla \Lambda(\mathbf{x})$ , with the gauge function

$$\Lambda(\mathbf{x}) = \int d^3x' \mathbf{A}_\perp(\mathbf{x}', t) \cdot \mathbf{K}_\perp(\mathbf{x}', \mathbf{x}). \quad (11)$$

The transverse part of the quantization function  $\mathbf{K}_\perp: (\mathbb{R}^3, \mathbb{R}^3) \rightarrow \mathbb{R}^3$  is nearly completely arbitrary [47] and uniquely fixes the gauge with respect to the fields. Choosing

$\chi_2$  to take the form in Eq. (9) allows one to write the entire vector potential in terms of its gauge-invariant transverse component, and a function which remains a  $c$ -number after quantization. All that remains is to specify the transverse part of the polarization, to describe a unique Lagrangian as per the prescription in Eq. (5). Noting that since we can write the longitudinal part as

$$\mathbf{P}_{\parallel}(\mathbf{x}, t) = - \int d^3x' \mathbf{K}_{\parallel}(\mathbf{x}, \mathbf{x}') \rho_A(\mathbf{x}', t), \quad (12)$$

we see that the additional term in the Lagrangian becomes (suppressing the time index)

$$- \frac{d}{dt} \int d^3x \left\{ \mathbf{A}_{\perp}(\mathbf{x}) \cdot \left[ \mathbf{P}_{\perp}(\mathbf{x}) + \int d^3x' \mathbf{K}_{\perp}(\mathbf{x}, \mathbf{x}') \rho_A(\mathbf{x}') \right] \right\}. \quad (13)$$

A convenient definition for the polarization is thus

$$\mathbf{P}(\mathbf{x}, t) = - \int d^3x' \mathbf{K}(\mathbf{x}, \mathbf{x}') \rho_A(\mathbf{x}', t), \quad (14)$$

such that the total time derivative vanishes, as

$$\int d^3x \mathbf{A}(\mathbf{x}, t) \cdot \mathbf{P}(\mathbf{x}, t) = 0, \quad (15)$$

and the polarization no longer appears at all in the Lagrangian; this allows one to quantize entirely on the basis of the gauge symmetry associated with the electromagnetic field (albeit in a subspace with the reduced gauge symmetry associated with the quantization constraint). Note that with this definition, a *gauge transformation* consists of the simultaneous change of the gauge function  $\mathbf{K}(\mathbf{x}, \mathbf{x}') \rightarrow \mathbf{K}'(\mathbf{x}, \mathbf{x}')$ , with an associated change in the vector potential and transverse polarization  $\mathbf{A}_{\parallel} \rightarrow \mathbf{A}'_{\parallel}$ ,  $\mathbf{P}_{\perp} \rightarrow \mathbf{P}'_{\perp}$  as determined by Eqs. (10) and (14). In Sec. V we extend this scheme to allow for *time-dependent* gauge conditions, by letting  $\mathbf{K}_{\perp}$  be an explicit function of time, such that a gauge transformation also changes  $\phi$ . The quantization function approach can also be used for quantization of systems with no explicit medium reservoir degrees of freedom, when dissipation and dispersion can be neglected in the dielectric function  $\epsilon$ , as we show in Appendix D.

Two important choices of gauge in the theory of light-matter interactions are the Coulomb and multipolar gauges; for the Coulomb gauge

$$\mathbf{K}_{\perp}^C(\mathbf{x}, \mathbf{x}') = 0, \quad (16)$$

and for the multipolar gauge,

$$\mathbf{K}_{\perp}^{\text{mp}}(\mathbf{x}, \mathbf{x}') = -(\mathbf{x}' - \mathbf{r}_A) \cdot \int_0^1 ds \delta^{\perp}[\mathbf{x} - \mathbf{r}_A - s(\mathbf{x}' - \mathbf{r}_A)], \quad (17)$$

where  $\mathbf{r}_A$  is an arbitrary  $c$ -number position, and  $\delta^{\perp} = \mathbf{I}\delta - \delta^{\parallel}$  is the transverse delta function (using dyadic notation with  $I_{ij} = \delta_{ij}$ ). This form is particularly useful for computing multipolar expansions for charge distributions localized around a position  $\mathbf{r}_A$  (e.g., a molecular center) and is closely related to the PZW transformation from the Coulomb to multipolar gauges, as discussed more in Sec. IV B.

## B. Canonical quantization with constraints

Having specified the gauge of the theory by means of the quantization function  $\mathbf{K}(\mathbf{x}, \mathbf{x}')$  (in particular, its transverse component), we can now apply Dirac's constrained quantization procedure using the constraints  $\chi_0 = \chi_1 = \chi_2 = 0$ . The details of this procedure are given in Appendix B. After quantization, we find that the theory can be expressed in terms of transverse canonical field variables  $\hat{\mathbf{A}}_{\perp}$  and  $\hat{\mathbf{\Pi}}$ , where

$$\hat{\mathbf{\Pi}} = \hat{\mathbf{\Pi}}_A - \hat{\mathbf{P}}, \quad (18)$$

as well as  $\hat{\mathbf{r}}_{\alpha}$ ,  $\hat{\mathbf{p}}_{\alpha}$ ,  $\hat{\mathbf{X}}_{\omega}$ , and  $\hat{\mathbf{\Pi}}_{\mathbf{x}_{\omega}}$ . These canonical coordinates satisfy the usual canonical commutation relations:

$$[\hat{\mathbf{A}}_{\perp}(\mathbf{x}), \hat{\mathbf{\Pi}}(\mathbf{x}')] = i\hbar \delta^{\perp}(\mathbf{x} - \mathbf{x}'), \quad (19a)$$

$$[\hat{\mathbf{r}}_{\alpha}, \hat{\mathbf{p}}_{\alpha'}] = i\hbar \mathbf{I} \delta_{\alpha\alpha'}, \quad (19b)$$

$$[\hat{\mathbf{X}}_{\omega}(\mathbf{x}), \hat{\mathbf{\Pi}}_{\omega'}(\mathbf{x}')] = i\hbar \mathbf{I} \delta(\mathbf{x} - \mathbf{x}') \delta(\omega - \omega'), \quad (19c)$$

with all other commutators vanishing.

## C. Fano diagonalization and Green's function expansion

As a result of the canonical commutation relations in Eqs. (19a)–(19c), the fields  $\hat{\mathbf{A}}_{\perp}$  and  $\hat{\mathbf{X}}_{\omega}$  as well as their conjugates  $\hat{\mathbf{\Pi}}$  and  $\hat{\mathbf{\Pi}}_{\mathbf{x}_{\omega}}$  can be expressed as a sum over bosonic creation and annihilation operators. Furthermore, the total Hamiltonian can be decomposed into a component which consists of only these operators,  $\hat{H}_F$ , as well as a component which involves the polarization operators associated with the  $\hat{\mathbf{r}}_{\alpha}$  operators and their conjugates  $\hat{H}_P$ :

$$\hat{H}_F = \frac{1}{2\epsilon_0} \int d^3x \left[ \hat{\mathbf{\Pi}} + \int d\omega \alpha(\mathbf{x}, \omega) \hat{\mathbf{X}}_{\omega} \right]^2 + \frac{1}{2\mu_0} (\nabla \times \hat{\mathbf{A}}_{\perp})^2 + \hat{H}_{\mathbf{x}_{\omega}} \quad (20)$$

with

$$\hat{H}_{\mathbf{x}_{\omega}} = \frac{1}{2} \int d^3x \int_0^{\infty} d\omega (\hat{\mathbf{\Pi}}_{\omega}^2 + \omega^2 \hat{\mathbf{X}}_{\omega}^2) \quad (21)$$

and

$$\hat{H}_P = \sum_{\alpha} \frac{[\hat{\mathbf{p}}_{\alpha} - q_{\alpha} \hat{\mathbf{A}}(\hat{\mathbf{r}}_{\alpha})]^2}{2m_{\alpha}} - \int d^3x \hat{\mathbf{P}}(\mathbf{x}) \cdot \hat{\mathbf{E}}(\mathbf{x}) - \frac{1}{2\epsilon_0} \int d^3x \hat{\mathbf{P}}^2(\mathbf{x}), \quad (22)$$

where Eq. (2b) was used to express the electric field operator in the quantization variables as

$$\hat{\mathbf{E}}(\mathbf{x}) = -\frac{1}{\epsilon_0} \left[ \hat{\mathbf{\Pi}}(\mathbf{x}) + \hat{\mathbf{P}}(\mathbf{x}) + \int_0^{\infty} d\omega \alpha(\mathbf{x}, \omega) \hat{\mathbf{X}}_{\omega}(\mathbf{x}) \right]. \quad (23)$$

The total Hamiltonian is then  $\hat{H} = \hat{H}_F + \hat{H}_P$ .

The term  $\hat{H}_F$  is quadratic in the quantization variables over bosonic fields, and as such, one should be able to, by a process of Fano diagonalization, express it in the form

$$\hat{H}_F = \int d^3x \int_0^{\infty} d\omega \hbar \omega \hat{\mathbf{b}}^{\dagger}(\mathbf{x}, \omega) \cdot \hat{\mathbf{b}}(\mathbf{x}, \omega), \quad (24)$$

where  $\hat{\mathbf{b}}(\mathbf{x}, \omega)$  are bosonic excitation operators which combine medium harmonic oscillator fields  $\hat{\mathbf{X}}_\omega$  and electromagnetic degrees of freedom, and satisfy

$$[\hat{\mathbf{b}}(\mathbf{x}, \omega), \hat{\mathbf{b}}^\dagger(\mathbf{x}', \omega')] = \mathbf{I}\delta(\mathbf{x} - \mathbf{x}')\delta(\omega - \omega'), \quad (25a)$$

$$[\hat{\mathbf{b}}(\mathbf{x}, \omega), \hat{\mathbf{b}}(\mathbf{x}', \omega')] = \mathbf{0}, \quad (25b)$$

$$[\hat{\mathbf{b}}^\dagger(\mathbf{x}, \omega), \hat{\mathbf{b}}^\dagger(\mathbf{x}', \omega')] = \mathbf{0}. \quad (25c)$$

For the Coulomb gauge, this is done precisely in the derivation by, e.g., Philbin [7]. Here we note that the results of this derivation can be applied directly with the substitution  $(\hat{\mathbf{A}}, \hat{\mathbf{\Pi}}_A) \rightarrow (\hat{\mathbf{A}}_\perp, \hat{\mathbf{\Pi}})$ . The important result is that the fields can be expressed as

$$\hat{\mathbf{E}}_F(\mathbf{x}) = i \int d^3x' \int_0^\infty d\omega \frac{1}{\epsilon_0\omega} \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) \cdot \hat{\mathbf{J}}_N(\mathbf{x}', \omega) + \text{H.c.}, \quad (26a)$$

$$\hat{\mathbf{A}}_\perp(\mathbf{x}) = \int d^3x' \int_0^\infty d\omega \frac{1}{\epsilon_0\omega^2} \mathbf{G}^\perp(\mathbf{x}, \mathbf{x}', \omega) \cdot \hat{\mathbf{J}}_N(\mathbf{x}', \omega) + \text{H.c.}, \quad (26b)$$

where

$$\hat{\mathbf{J}}_N(\mathbf{x}, \omega) = \sqrt{\frac{\hbar\omega}{2}} \alpha(\mathbf{x}, \omega) \hat{\mathbf{b}}(\mathbf{x}, \omega), \quad (27)$$

and  $\hat{\mathbf{E}}_F$  is the part of the total electric field operator  $\hat{\mathbf{E}} = \hat{\mathbf{E}}_F - \hat{\mathbf{P}}/\epsilon_0$  which can be expressed in terms of the bosonic operators that diagonalize  $\hat{H}_F$ . In accordance with our assumption of nonmagnetic media, the photonic Green's function  $\mathbf{G}(\mathbf{x}, \mathbf{x}', \omega)$  is a tensor (or dyad) which satisfies the Helmholtz equation for a dipole source:

$$\left[ \nabla^x \times \nabla^x \times - \frac{\omega^2}{c^2} \epsilon(\mathbf{x}, \omega) \right] \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) = \frac{\omega^2}{c^2} \mathbf{I}\delta(\mathbf{x} - \mathbf{x}'), \quad (28)$$

together with the corresponding retarded boundary conditions. For example, for open dielectrics (i.e., cavity resonators surrounded by a homogeneous medium with index of refraction  $n_B$ ), one can use the Silver-Müller radiation condition:

$$\frac{\mathbf{x}}{|\mathbf{x}|} \times \nabla^x \times \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega) \rightarrow in_B \frac{\omega}{c} \mathbf{G}(\mathbf{x}, \mathbf{x}', \omega), \quad (29)$$

which holds as  $|\mathbf{x}| \rightarrow \infty$ . Here we let the notation  $\mathbf{G}^\perp(\mathbf{x}, \mathbf{x}', \omega)$  refer to the transverse part of  $\mathbf{G}(\mathbf{x}, \mathbf{x}', \omega)$  with respect to the left-hand side of the dyad, and spatial argument  $\mathbf{x}$ .

#### IV. GAUGE INVARIANCE AND HILBERT SPACE TRUNCATION

In this section we first show in Sec. IV A how gauge invariance manifests in the quantized theory, and how gauge transformations can be implemented as unitary transformations within the general gauge function quantization method. We then discuss how material and mode truncation can potentially break this gauge invariance, and how this can be avoided, in Secs. IV B and IV C. Our main contribution in this section is the arbitrary-gauge Hamiltonian under material and mode truncation  $\hat{H}_L$ , given by Eq. (69), from which we

resolve gauge ambiguities and derive simplified models in later sections.

##### A. Gauge invariance in the quantum theory

Prior to quantization, gauge invariance manifests as the invariance of the Lagrangian under gauge transformations of the four-potential. After quantization using the arbitrary gauge approach [in terms of the quantization function  $\mathbf{K}(\mathbf{x}, \mathbf{x}')$ ], however, the hallmark of gauge invariance is the invariance of the Schrödinger equation under a local phase change of the state vector simultaneous with a gauge transformation of the potentials. Such a process can be implemented as a unitary transformation which transforms the system from one *fixed* gauge to another [51].

As mentioned in the previous section, a gauge transformation consists of the change  $\mathbf{K}(\mathbf{x}, \mathbf{x}') \rightarrow \mathbf{K}'(\mathbf{x}, \mathbf{x}')$  and the associated change in the longitudinal component of the vector potential and the transverse component of the polarization. Thus, for the theory to be gauge invariant, this change should be compensated by a local phase change in the state vector, implemented as a unitary operator.

Specifically, for the quantization scheme developed here and the set of gauge transformations allowed therein, this phase variation can be expressed as a unitary transformation  $|\psi\rangle \rightarrow \hat{W}|\psi\rangle$ , with an accompanying change of the Hamiltonian  $\hat{W}\hat{H}\hat{W}^\dagger$ , in order to preserve the form of the Schrödinger equation evolution. To be concrete, consider two gauges indexed by  $g$  and  $g'$ , with quantization functions  $\mathbf{K}^g(\mathbf{x}, \mathbf{x}')$  and  $\mathbf{K}^{g'}(\mathbf{x}, \mathbf{x}')$ , respectively. Then a gauge transformation from gauge representation  $g$  to  $g'$  can be found as

$$\begin{aligned} \hat{W}_{g'g} &= \exp \left\{ \frac{i}{\hbar} \sum_\alpha q_\alpha [\hat{\Lambda}_{g'}(\hat{\mathbf{r}}_\alpha) - \hat{\Lambda}_g(\hat{\mathbf{r}}_\alpha)] \right\} \\ &= \exp \left\{ \frac{i}{\hbar} \int d^3x [\hat{\Lambda}_{g'}(\mathbf{x}) - \hat{\Lambda}_g(\mathbf{x})] \hat{\rho}_A(\mathbf{x}) \right\} \\ &= \exp \left\{ -\frac{i}{\hbar} \int d^3x [\hat{\mathbf{P}}_\perp^{g'}(\mathbf{x}) - \hat{\mathbf{P}}_\perp^g(\mathbf{x})] \cdot \hat{\mathbf{A}}_\perp(\mathbf{x}) \right\}, \quad (30) \end{aligned}$$

and  $\hat{\Lambda}_g(\mathbf{x})$  is the quantized version of the gauge function expressed in Eq. (11) for a gauge indexed by  $g$ . The state vector transforms as  $|\psi_{g'}\rangle = \hat{W}_{g'g}|\psi_g\rangle$ . To determine the transformation effect on the variables constituting the Hamiltonian, it is useful to note the following relations (suppressing frequency indices):

$$\begin{aligned} &[\hat{A}_{\perp,i}(\mathbf{x}), \hat{A}_{\perp,j}(\mathbf{x}')] \\ &= \frac{\hbar}{2\epsilon_0^2} \int d^3r \int_0^\infty d\omega \frac{\alpha^2(\mathbf{r})}{\omega^3} G_{ik}^\perp(\mathbf{x}, \mathbf{r}) G_{kj}^{\perp*}(\mathbf{r}, \mathbf{x}') - \text{H.c.} \\ &= \frac{\hbar}{\pi\epsilon_0} \int_0^\infty \frac{d\omega}{\omega^2} \text{Im}[G_{ij}^\perp(\mathbf{x}, \mathbf{x}')] - \text{H.c.} \\ &= 0, \quad (31) \end{aligned}$$

where we have used the Green's function relations [4]

$$G_{ij}(\mathbf{x}, \mathbf{x}', \omega) = G_{ji}(\mathbf{x}', \mathbf{x}, \omega) \quad (32)$$

and

$$\int d^3r \epsilon_I(\mathbf{r}, \omega) \mathbf{G}(\mathbf{x}, \mathbf{r}, \omega) \cdot \mathbf{G}(\mathbf{r}, \mathbf{x}', \omega) = \text{Im}[\mathbf{G}(\mathbf{x}, \mathbf{x}', \omega)]. \quad (33)$$

Equations (31) and (10) together imply that  $[\hat{A}_i(\mathbf{x}), \hat{A}_j(\mathbf{x}')] = 0$  as well.

Of particular use is the transformation of the canonical momenta:

$$\hat{W}_{g'g}[\hat{\Pi}(\mathbf{x}) + \hat{\mathbf{P}}^g(\mathbf{x})]\hat{W}_{g'g}^\dagger = \hat{\Pi}(\mathbf{x}) + \hat{\mathbf{P}}^{g'}(\mathbf{x}), \quad (34)$$

$$\hat{W}_{g'g}[\hat{\mathbf{p}}_\alpha - q_\alpha \hat{\mathbf{A}}^g(\hat{\mathbf{r}}_\alpha)]\hat{W}_{g'g}^\dagger = \hat{\mathbf{p}}_\alpha - q_\alpha \hat{\mathbf{A}}^{g'}(\hat{\mathbf{r}}_\alpha). \quad (35)$$

To show the effect of this transformation on the Hamiltonian, it is useful to write the arbitrary gauge Hamiltonian  $\hat{H}^g$  in the form

$$\begin{aligned} \hat{H}^g = & \int d^3x \left[ \frac{\epsilon_0}{2} [\hat{\mathbf{E}}^g(\mathbf{x})]^2 + \frac{1}{2\mu_0} \hat{\mathbf{B}}^2(\mathbf{x}) \right] \\ & + \hat{H}_{\mathbf{x}_\omega} + \sum_\alpha \frac{[\hat{\mathbf{p}}_\alpha - q_\alpha \hat{\mathbf{A}}^g(\hat{\mathbf{r}}_\alpha)]^2}{2m_\alpha}, \end{aligned} \quad (36)$$

where

$$\hat{\mathbf{E}}^g(\mathbf{x}) = -\frac{1}{\epsilon_0} \left[ \hat{\Pi}(\mathbf{x}) + \hat{\mathbf{P}}^g(\mathbf{x}) + \int_0^\infty d\omega \alpha(\mathbf{x}, \omega) \hat{\mathbf{X}}_\omega(\mathbf{x}) \right]. \quad (37)$$

The unitary transformation has no effect on  $\hat{\mathbf{B}}$  [as a result of Eq. (31)] or the reservoir operators  $\hat{\mathbf{X}}_\omega$ ,  $\hat{\Pi}_{\mathbf{x}_\omega}$ , and clearly  $\hat{W}_{g'g} \hat{\mathbf{E}}^g(\mathbf{x}) \hat{W}_{g'g}^\dagger = \hat{\mathbf{E}}^{g'}(\mathbf{x})$ , as a consequence of Eq. (34). Using this fact and Eq. (35), it is easy to see that the unitary transform has the effect of performing the gauge transform, and thus we confirm the quantum theory is indeed gauge invariant.

## B. Material truncation

The issue with truncation is that the transformation given by Eq. (35) necessarily requires the full infinite dimensional operator algebra to be implemented, and thus any method of truncating the operator  $\hat{W}_{g'g}$  (which must be truncated to operate on the reduced dimensionality state vector) will fail to give the necessary gauge transformation [22,23,25,26,28,68]. This results in a theory which does not respect the gauge principle and gives ambiguous results, especially in the USC regime. This can also be seen as the truncation creating a nonlocal potential which can be expressed as a function of truncated momentum operators, to which the minimal coupling replacement that ensures gauge invariance has not been applied [25,27,69] or inconsistent constraining of interactions to a specific subspace [29].

To circumvent this, we can instead write the Hamiltonian, prior to truncation, in terms of unitary operators which implement the minimal coupling replacement on a Hamiltonian which, in the absence of transverse coupling to the electric field (e.g., the eigenstates of a molecular system), has a discrete set of energy levels which are near-resonant with relevant medium-assisted interactions with the electric field. The truncation can then be applied directly to the position operators, which ensures the gauge transformation is consistent with the reduced Hilbert space dimensionality and preserves

gauge invariance [27,29,70]. This procedure is also consistent with a lattice gauge theory perspective, where the local phase transformation acts on the state vector only at discrete ‘‘lattice’’ points in space equal in number to the number of states left after truncation [31], as well as the Peierls substitution for introducing electromagnetic interactions within tight-binding models [70].

To identify the ‘‘bare’’ matter Hamiltonian, containing only the interparticle Coulomb interactions between the constituent particles, note that in the absence of coupling with the transverse field, the material Hamiltonian becomes

$$\hat{H}_0 = \sum_\alpha \frac{\hat{\mathbf{p}}_\alpha^2}{2m_\alpha} + \sum_{\alpha, \alpha'} \hat{V}_{\text{Coul}}(\hat{\mathbf{r}}_\alpha, \hat{\mathbf{r}}_{\alpha'}), \quad (38)$$

where

$$\begin{aligned} \sum_{\alpha, \alpha'} \hat{V}_{\text{Coul}}(\hat{\mathbf{r}}_\alpha, \hat{\mathbf{r}}_{\alpha'}) &= \int d^3x \frac{\hat{\mathbf{P}}_\parallel^2(\mathbf{x})}{2\epsilon_0} \\ &= \frac{1}{2} \sum_{\alpha, \alpha'} \frac{q_\alpha q_{\alpha'}}{4\pi \epsilon_0 |\hat{\mathbf{r}}_\alpha - \hat{\mathbf{r}}_{\alpha'}|}, \end{aligned} \quad (39)$$

and we are assuming that the medium-assisted longitudinal field is sufficiently weak that it suffices to use the unscreened Coulomb potential to calculate the unperturbed material eigenstates—alternatively, we can simply phenomenologically use the eigenstates which are corrected by the medium-assisted longitudinal field as the basis  $\hat{H}_0$  (see Appendix D and Ref. [65] for an analogous discussion in the case of a nondispersive and nonabsorbing medium).

In this manner, we can write the entire arbitrary-gauge Hamiltonian as (for a gauge indexed by ‘‘g’’),

$$\hat{H}^g = \hat{H}_F + \hat{H}_0 + \hat{H}_{\text{int}}^g, \quad (40)$$

where

$$\begin{aligned} \hat{H}_{\text{int}}^g = & - \sum_\alpha \frac{q_\alpha}{m_\alpha} \hat{\mathbf{p}}_\alpha \cdot \hat{\mathbf{A}}^g(\hat{\mathbf{r}}_\alpha) + \sum_\alpha \frac{q_\alpha^2}{2m_\alpha} [\hat{\mathbf{A}}^g(\hat{\mathbf{r}}_\alpha)]^2 \\ & - \int d^3x \hat{\mathbf{P}}(\mathbf{x}) \cdot \hat{\mathbf{E}}_F^g(\mathbf{x}) + \frac{1}{2\epsilon_0} \int d^3x [\hat{\mathbf{P}}_\perp^g(\mathbf{x})]^2. \end{aligned} \quad (41)$$

Next, we introduce the unitary operator

$$\hat{U}_g = \exp \left[ \frac{i}{\hbar} \int d^3x \hat{\mathbf{A}}^g(\mathbf{x}) \cdot \hat{\mathbf{Z}}_A(\mathbf{x}) \right], \quad (42)$$

where

$$\hat{\mathbf{Z}}_A(\mathbf{x}) = \sum_\alpha q_\alpha \hat{\mathbf{r}}_\alpha \int_0^1 ds \delta(\mathbf{x} - s\hat{\mathbf{r}}_\alpha) \quad (43)$$

is an operator chosen to implement, approximately, the minimal coupling transformation  $\hat{\mathbf{p}}_\alpha \rightarrow \hat{\mathbf{p}}_\alpha - q_\alpha \hat{\mathbf{A}}^g(\hat{\mathbf{r}}_\alpha)$ , as appears in the  $\hat{H}_F$  part of the full Hamiltonian  $\hat{H}^g$  [Eq. (20)]. In actuality, the full transformation of the particle momenta under  $\hat{U}_g$  is

$$\hat{U}_g \hat{\mathbf{p}}_\alpha \hat{U}_g^\dagger = \hat{\mathbf{p}}_\alpha - q_\alpha \hat{\mathbf{A}}^g(\hat{\mathbf{r}}_\alpha) - q_\alpha \hat{\mathbf{r}}_\alpha \times \int_0^1 ds s \hat{\mathbf{B}}(s\hat{\mathbf{r}}_\alpha). \quad (44)$$

Note that we also have  $\hat{U}_g = \hat{W}_{g,\text{mp}}$ ; this transformation takes an operator from the multipolar gauge to gauge  $g$ . We give this operator as it appears in this context its own symbol  $\hat{U}_g$  to emphasize its role in restoring gauge invariance under material truncation, described in the following.

Using the transformation of Eq. (44), we can write the full Hamiltonian as

$$\hat{H}^g = \hat{H}_F + \hat{U}_g \hat{H}_0 \hat{U}_g^\dagger - \int d^3x \hat{\mathbf{P}}(\mathbf{x}) \cdot \hat{\mathbf{E}}_F^g(\mathbf{x}) + \frac{1}{2\epsilon_0} \int d^3x [\hat{\mathbf{P}}_\perp^g(\mathbf{x})]^2 + \hat{H}_{\text{mag}}^g, \quad (45)$$

where

$$\hat{H}_{\text{mag}}^g = \sum_\alpha \frac{q_\alpha}{2m_\alpha} \int_0^1 ds s \left\{ \hat{\mathbf{p}}_\alpha \cdot \hat{\mathbf{r}}_\alpha \times \hat{\mathbf{B}}(s\hat{\mathbf{r}}_\alpha) + \hat{\mathbf{r}}_\alpha \times \hat{\mathbf{B}}(s\hat{\mathbf{r}}_\alpha) \cdot \hat{\mathbf{p}}_\alpha - 2q_\alpha \hat{\mathbf{A}}^g(\hat{\mathbf{r}}_\alpha) \cdot \hat{\mathbf{r}}_\alpha \times \hat{\mathbf{B}}(s\hat{\mathbf{r}}_\alpha) - q_\alpha \int_0^1 ds' s' [\hat{\mathbf{r}}_\alpha \times \hat{\mathbf{B}}(s'\hat{\mathbf{r}}_\alpha)] \cdot [\hat{\mathbf{r}}_\alpha \times \hat{\mathbf{B}}(s'\hat{\mathbf{r}}_\alpha)] \right\}. \quad (46)$$

The magnetic terms were analyzed in detail in Ref. [18] (in that case, as they appear in the multipolar gauge), with the conclusion that, for atomic systems, the scaling of these terms relative to the electric dipole interaction is proportional to  $\sim (Z_{\text{eff}}\alpha_0)^2$ , where  $Z_{\text{eff}}$  is the effective (screened) charge of the nucleus, and  $\alpha_0$  is the fine-structure constant. For nonrelativistic systems,  $Z_{\text{eff}}\alpha_0 \ll 1$ , and so going forward we will neglect the influence of  $\hat{H}_{\text{mag}}^g$ . We note that this is an inherent approximation of the theory (although one that is very well founded in most circumstances), which has to our knowledge has not been acknowledged to date in the literature on restoring gauge invariance in truncated material systems.

The Hamiltonian of Eq. (45), after neglecting the magnetic terms, is gauge-invariant under material truncation *provided* we truncate the position operators  $\hat{\mathbf{x}}_\alpha \equiv \hat{P}\hat{\mathbf{r}}_\alpha\hat{P}$ , where

$$\hat{P} = \sum_{i=1}^N |\phi_i\rangle \langle \phi_i| \quad (47)$$

is a projector operator onto a finite set of  $N$  eigenstates of  $\hat{H}_0$ , such that in the truncated space, we have

$$\hat{P}\hat{H}_0\hat{P} = \hat{H}_0 = \sum_{i=1}^N \hbar\omega_i |\phi_i\rangle \langle \phi_i|, \quad (48)$$

where  $|\phi_i\rangle$  denotes the  $i$ th eigenstate of  $\hat{H}_0$  with energy  $\hbar\omega_i$ . We use throughout this work calligraphic characters to denote operators which act on the truncated space. We stress that, except for  $\hat{H}_0$ , the (correctly) truncated operators are those that are expressed in terms of the projected position operators  $\hat{\mathbf{x}}_\alpha$ , and *not* those with the projector operator  $\hat{P}$  directly applied, which generally will violate gauge invariance.

We then take  $\hat{\mathbf{P}}^g \rightarrow \hat{\mathcal{P}}^g$ , where

$$\hat{\mathcal{P}}^g = - \sum_\alpha q_\alpha \mathbf{K}^g(\mathbf{x}, \hat{\mathbf{x}}_\alpha) \quad (49)$$

and  $\hat{U}_g \rightarrow \hat{\mathcal{U}}_g$ :

$$\hat{\mathcal{U}}_g = \exp \left[ \frac{i}{\hbar} \int d^3x \hat{\mathbf{A}}^g(\mathbf{x}) \cdot \hat{\mathcal{Z}}_A(\mathbf{x}) \right] = \exp \left[ \frac{i}{\hbar} \sum_\alpha q_\alpha \hat{\mathbf{x}}_\alpha \cdot \int_0^1 ds \hat{\mathbf{A}}^g(s\hat{\mathbf{x}}_\alpha) \right], \quad (50)$$

where  $\hat{\mathcal{Z}}_A$  is  $\hat{\mathbf{Z}}_A$  expressed in terms of the truncated position operators  $\hat{\mathbf{x}}_\alpha$ . Thus, the truncated material basis arbitrary-gauge Hamiltonian  $\hat{\mathcal{H}}^g$  can be written as

$$\hat{\mathcal{H}}^g = \hat{H}_F + \hat{\mathcal{U}}_g \hat{H}_0 \hat{\mathcal{U}}_g^\dagger - \int d^3x \hat{\mathcal{P}}^g(\mathbf{x}) \cdot \hat{\mathbf{E}}_F^g(\mathbf{x}) + \frac{1}{2\epsilon_0} \int d^3x [\hat{\mathcal{P}}_\perp^g(\mathbf{x})]^2. \quad (51)$$

In the Coulomb gauge,  $\mathbf{K}_\perp^C = \hat{\mathbf{P}}_\perp^C = \hat{\mathbf{A}}_\parallel^C = 0$ , and the Hamiltonian,  $\hat{\mathcal{H}}^C$ , can be expressed as

$$\hat{\mathcal{H}}^C = \hat{H}_F + \hat{\mathcal{U}}_C \hat{H}_0 \hat{\mathcal{U}}_C^\dagger - \int d^3x \hat{\mathcal{P}}_\parallel(\mathbf{x}) \cdot [\hat{\mathbf{E}}_F(\mathbf{x})]_\parallel. \quad (52)$$

In the multipolar gauge,  $\mathbf{K}^{\text{mp}}$  is given by Eq. (17), which implies (again taking  $\mathbf{r}_A = 0$ )

$$\hat{\mathcal{P}}_\perp^{\text{mp}}(\mathbf{x}) = \sum_\alpha q_\alpha \hat{\mathbf{x}}_\alpha \cdot \int_0^1 ds \delta^\perp(\mathbf{x} - s\hat{\mathbf{x}}_\alpha) \quad (53)$$

and

$$\hat{\mathbf{A}}^{\text{mp}}(\mathbf{x}) = - \int_0^1 ds \mathbf{x} \times \hat{\mathbf{B}}(s\mathbf{x}). \quad (54)$$

In the multipolar gauge,  $\mathbf{x} \cdot \hat{\mathbf{A}}^{\text{mp}}(\mathbf{x}) = 0$ , and from this it is easy to show that  $\hat{\mathcal{U}}_{\text{mp}} = 1$ . Applying this result, we thus find

$$\hat{\mathcal{H}}^{\text{mp}} = \hat{H}_F + \hat{\mathcal{H}}_0 - \sum_\alpha q_\alpha \hat{\mathbf{x}}_\alpha \cdot \int_0^1 ds \hat{\mathbf{E}}_F^{\text{mp}}(s\hat{\mathbf{x}}_\alpha) + \sum_{\alpha,\alpha'} \frac{q_\alpha q_{\alpha'}}{2\epsilon_0} \int_0^1 ds \int_0^1 ds' \hat{\mathbf{x}}_\alpha \cdot \delta^\perp(s'\hat{\mathbf{x}}_{\alpha'} - s\hat{\mathbf{x}}_\alpha) \cdot \hat{\mathbf{x}}_{\alpha'}. \quad (55)$$

As in the untruncated theory, we can implement a gauge change from one fixed gauge to another by means of a unitary transformation:

$$\hat{\mathcal{W}}_{g'g} \hat{\mathcal{H}}^g \hat{\mathcal{W}}_{g'g}^\dagger, \quad (56)$$

where  $\hat{\mathcal{W}}_{g'g}$ , defined from Eq. (30), is expressed in terms of the truncated position operators  $\hat{\mathbf{x}}_\alpha$ :

$$\hat{\mathcal{W}}_{g'g} = \exp \left[ - \frac{i}{\hbar} \int d^3x [\hat{\mathcal{P}}_\perp^{g'}(\mathbf{x}) - \hat{\mathcal{P}}_\perp^g(\mathbf{x})] \cdot \hat{\mathbf{A}}_\perp(\mathbf{x}) \right]. \quad (57)$$

It is straightforward to verify that the transformation in Eq. (56) is equivalent to replacing  $\hat{\mathbf{A}}^g$  and  $\hat{\mathcal{P}}^g$  in the arbitrary-gauge Hamiltonian  $\hat{\mathcal{H}}$  in Eq. (51) with  $\hat{\mathbf{A}}^{g'}$  and  $\hat{\mathcal{P}}^{g'}$  (or equivalently, replacing  $\mathbf{K}^g$  with  $\mathbf{K}^{g'}$ ), respectively; thus we see that gauge invariance is preserved under material truncation in the fully quantized theory.

It is worth noting that  $\hat{\mathbf{Z}}_A$  is precisely the multipolar polarization  $\hat{\mathbf{P}}^{\text{mp}}(\mathbf{x})$ , and as such,  $\hat{U}_g = \hat{W}_{g,\text{mp}}$ , as previously noted. Moreover, if it is evaluated in the Coulomb gauge, then  $\hat{U}_C$

is the unitary operator which implements the well-known PZW transformation [45,46,50]. The PZW transformation  $\hat{U}_C^\dagger \hat{H}^C \hat{U}_C = \hat{H}^{\text{mp}}$  removes the transformation that generates minimal coupling from  $\hat{H}_0$ , which allows one to truncate the energy levels of the bare system without needing to rely on the infinite dimensional operator algebra required to transform the  $\hat{\mathbf{p}}_\alpha$  operators. This is why naive truncation (in the sense of  $\hat{H} \rightarrow \hat{P}\hat{H}\hat{P}$ ) in the multipolar gauge gives much more accurate results than the Coulomb gauge [23] and is in fact generally assumed to not break gauge invariance [53]. It is worth noting that this argument relies on the neglect of the magnetic terms, however, and should be understood as a nonrelativistic approximation. As discussed in the following section, naive truncation in the multipolar gauge also fails in general when anything less than a complete set of modes is used to expand the electromagnetic fields, which is often the case in, for example, cavity QED.

### C. Mode truncation

In addition to material truncation, we can also consider truncation of the transverse electrodynamic degrees of freedom: for example, a mode truncation, where the “modes” are typically solutions to Maxwell’s equations subject to a certain boundary condition (e.g., fixed, periodic, or open). To do so, note that the arbitrary gauge Hamiltonian from Eq. (51) can also be written as

$$\hat{\mathcal{H}}^s = \hat{V}_g \hat{H}_F \hat{V}_g^\dagger + \hat{U}_g \hat{\mathcal{H}}_0 \hat{U}_g^\dagger - \int d^3x [\hat{\mathbf{E}}_F(\mathbf{x})]_{\parallel} \cdot \hat{\mathbf{P}}_{\parallel}(\mathbf{x}), \quad (58)$$

where

$$\hat{V}_g = \exp \left[ -\frac{i}{\hbar} \int d^3x \hat{\mathbf{P}}^s(\mathbf{x}) \cdot \hat{\mathbf{A}}_{\perp}(\mathbf{x}) \right]. \quad (59)$$

Note that  $\hat{V}_g = \hat{W}_{g,C}$ ; this transformation takes an operator from its Coulomb gauge representation to a generic one.

Similar to the analysis in the case of material truncation, one can show that the transformation induced by  $\hat{V}_g$  relies on the operator relationship  $[\hat{\mathbf{A}}_{\perp}(\mathbf{x}), \hat{\mathbf{\Pi}}(\mathbf{x}')] = i\hbar \delta^{\perp}(\mathbf{x} - \mathbf{x}')$ , which requires a *complete set of transverse modes* to expand the photonic operators in. As such, if the number of modes included in the system Hamiltonian is to be truncated naively, a gauge transformation in the reduced space can no longer be implemented as a unitary evolution, violating gauge invariance. Equivalently, this can be seen as not properly introducing coupling between the truncated subspaces of the system consistently [24]. Truncation of the Fock space photon number also breaks gauge invariance in this manner, although we focus on the case of mode truncation in this work.

To be explicit, consider a mode projection operator  $\hat{P}_M$  that satisfies

$$\hat{P}_M \hat{\mathbf{A}}_{\perp}(\mathbf{x}) \hat{P}_M = \hat{\mathbf{A}}_{\perp}(\mathbf{x}). \quad (60)$$

The Hamiltonian which retains gauge invariance under mode truncation is then simply Eq. (58), but with  $\hat{V}_g \rightarrow \hat{V}_g$ , and  $\hat{\mathcal{H}}_0$  is  $\hat{\mathcal{H}}_0$  evaluated in terms of  $\hat{\mathbf{A}}_{\perp}$  instead of  $\hat{\mathbf{A}}_{\perp}$ .

To give a concrete example, let us consider a modal expansion for the transverse vector potential:

$$\hat{\mathbf{A}}_{\perp}(\mathbf{x}) = \sum_{\mu} \sqrt{\frac{\hbar}{2\epsilon_0 \chi_{\mu\mu}}} \mathbf{f}_{\mu}(\mathbf{x}) \hat{a}_{\mu} + \text{H.c.}, \quad (61)$$

where we denote the mode expansion over a *finite* sum of “relevant” modes with transverse mode profiles  $\mathbf{f}_{\mu}(\mathbf{x})$  and annihilation (creation) operators  $\hat{a}_{\mu}$  ( $\hat{a}_{\mu}^\dagger$ ). We can then define the projection operator as

$$\hat{P}_M = \bigotimes_{\mu} \sum_{n_{\mu}=0}^{\infty} |n_{\mu}\rangle \langle n_{\mu}|. \quad (62)$$

Subsequently, the bosonic Hamiltonian is

$$\begin{aligned} \hat{\mathcal{H}}_F &= \int d^3x \int_0^{\infty} d\omega \omega \hat{P}_M \hat{\mathbf{b}}^\dagger(\mathbf{x}, \omega) \hat{P}_M \cdot \hat{P}_M \hat{\mathbf{b}}(\mathbf{x}, \omega) \hat{P}_M \\ &= \sum_{\mu\nu} \hbar \chi_{\mu\nu} \hat{a}_{\mu}^\dagger \hat{a}_{\nu}. \end{aligned} \quad (63)$$

In Appendix C we give more details on the construction of these discrete modes from the continuum, and their relationship to the Hermitian matrix  $\chi_{\mu\nu}$ . It is important to note that the mode functions  $\mathbf{f}_{\mu}$  are not the usual normal mode solutions to the Helmholtz equation, but rather nonorthogonal transverse modal expansion functions which satisfy, if the truncation is not applied,

$$\sum_{\mu\nu} \frac{\chi_{\mu\nu}}{\sqrt{\chi_{\mu\mu} \chi_{\nu\nu}}} \mathbf{f}_{\mu}(\mathbf{x}) \mathbf{f}_{\nu}^*(\mathbf{x}') = \delta^{\perp}(\mathbf{x} - \mathbf{x}'), \quad (64)$$

as shown in Appendix C.

Note that while throughout we refer to these as “modes,” they are, more generally, a truncation of the spatial and frequency-dependent degrees of freedom of the electromagnetic fields and passive medium reservoir fields. Specifically, the truncation process involves a projection of a spatial and frequency-dependent orthonormal basis onto the polariton operators  $\hat{\mathbf{b}}(\mathbf{x}, \omega)$ ,  $\hat{\mathbf{b}}^\dagger(\mathbf{x}, \omega)$ , and need not necessarily satisfy the Helmholtz equation with appropriate boundary conditions—although for truncation to be a useful approximation technique, this is presumed to be the case. A consequence of this is that  $\hat{\mathcal{H}}_F$  is not diagonal with respect to the finite mode basis (an effect known from, e.g., quantized QNMs [39,41] and quasimodes [34], as well as supermodes in quantum nonlinear optics [71]). This is why, in contrast to the case of material truncation, we define the mode truncation with respect to the field expansion itself, and not the field Hamiltonian; in the case of a nondispersive and nonabsorbing medium, it is possible to truncate with respect to the true normal modes of the medium, and both approaches are then equivalent. In Appendix D we discuss this case in more detail.

As an important example, in Appendix C we show how the discrete modes can be chosen to correspond over a restricted region of space to QNMs—although for this case the completeness relation (64) does not apply directly, as the expansion is valid only over a spatial region where the QNMs form a well-behaved basis for the transverse Green’s function. Also note that even in a dielectric medium with permittivity that is real and independent of frequency, it is often useful to

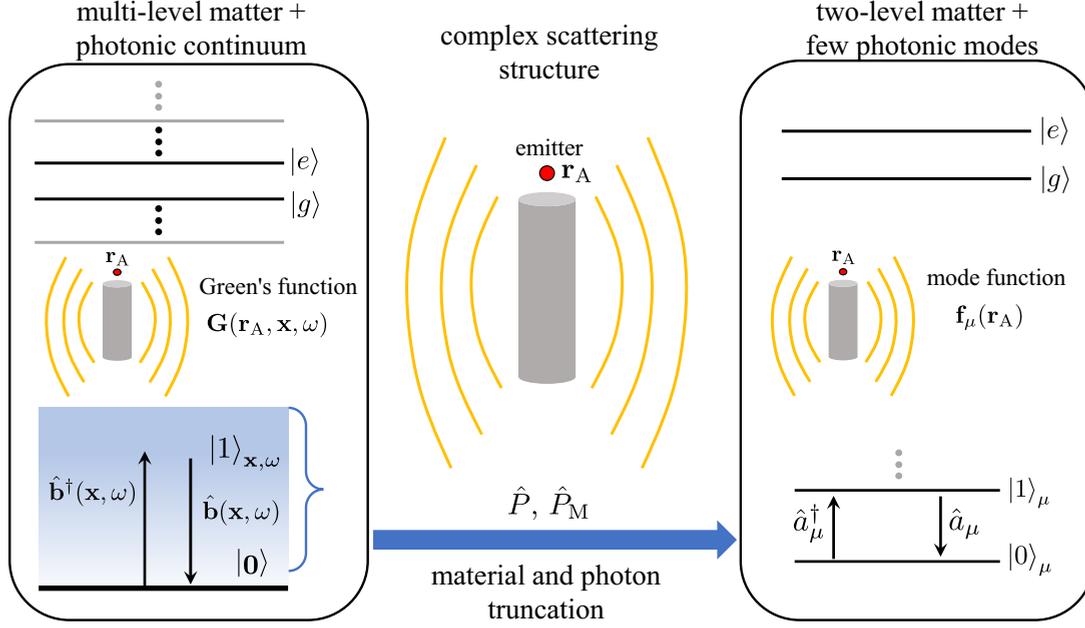


FIG. 2. Visualization of the photon and material truncation. An exemplary system consisting of a dielectric rod and a pointlike emitter can generally be described quantum mechanically by a multilevel electronic system interacting with a continuum of photon modes (left), which reflect hybridized states of the electromagnetic fields and the passive dielectric material. The electromagnetic field depends on a three-dimensional integration over the photon Green's function. The proper truncation of the material and photon degrees of freedom allows for a description of the quantum system with few energy levels (right), while still preserving gauge invariance. The electromagnetic field is obtained from the optical mode functions.

use mode expansions which are not the exact “true modes” of the entire system. For example, quasimodes [34], which use an artificial permittivity to obtain mode functions which represent an idealized version of the system of interest, and QNMs, where the open-boundary conditions lead to non-Hermitian eigenvalues even without dispersion or absorption.

Applying the correctly truncated unitary transform to the bosonic Hamiltonian,  $\hat{\mathcal{V}}_g \hat{a}_\mu \hat{\mathcal{V}}_g^\dagger = \hat{a}_\mu - i \hat{\xi}_\mu^g$ , we obtain

$$\hat{\mathcal{V}}_g \hat{\mathcal{H}}_F \hat{\mathcal{V}}_g^\dagger = \hat{\mathcal{H}}_F + \left( i \sum_{\mu\nu} \hbar \chi_{\mu\nu}^* \hat{a}_\mu \hat{\xi}_\nu^{g\dagger} + \text{H.c.} \right) + \sum_{\mu\nu} \hbar \chi_{\mu\nu} \hat{\xi}_\mu^{g\dagger} \hat{\xi}_\nu^g, \quad (65)$$

where

$$\hat{\xi}_\mu^g = \sum_\alpha \frac{q_\alpha}{\sqrt{2\epsilon_0 \hbar \chi_{\mu\mu}}} \int d^3x \mathbf{K}_\perp^g(\mathbf{x}, \hat{\mathbf{z}}_\alpha) \cdot \mathbf{f}_\mu^*(\mathbf{x}). \quad (66)$$

The second term in Eq. (65) gives the transverse coupling between photonic and material subspaces, and can also be written as

$$\left( i \sum_{\mu\nu} \hbar \chi_{\mu\nu}^* \hat{a}_\mu \hat{\xi}_\nu^{g\dagger} + \text{H.c.} \right) = - \int d^3x [\hat{\mathcal{E}}_F(\mathbf{x})]_\perp \cdot \hat{\mathcal{P}}_\perp^g(\mathbf{x}), \quad (67)$$

where  $[\hat{\mathcal{E}}_F]_\perp$  is the part of the *correctly mode-truncated* transverse electric field operator that can be expressed in terms of

the bosonic operators:

$$[\hat{\mathcal{E}}_F(\mathbf{x})]_\perp = i \sum_\mu \sqrt{\frac{\hbar \chi_{\mu\mu}}{2\epsilon_0}} \mathbf{f}'_\mu(\mathbf{x}) \hat{a}_\mu + \text{H.c.}, \quad (68)$$

and  $\mathbf{f}'_\mu = \sum_\nu \frac{\chi_{\mu\nu}^*}{\sqrt{\chi_{\mu\mu} \chi_{\nu\nu}}} \mathbf{f}_\nu$ . The quantum system after material and photon truncation is visualized in Fig. 2.

In a general gauge, the full correctly truncated transverse electric field operator can be found from  $\hat{\mathcal{E}}_\perp^g = \hat{\mathcal{V}}_g [\hat{\mathcal{E}}_F]_\perp \hat{\mathcal{V}}_g^\dagger$ . One can show that this is the correctly truncated form of the transverse electric field operator by enforcing  $\hat{\mathcal{E}}_\perp^g = -\frac{\partial}{\partial t} \hat{\mathcal{A}}_\perp$ , and applying the Heisenberg equation of motion. Of course, we could also define a mode expansion initially with respect to the transverse electric field, however, this would violate gauge invariance under mode truncation and not properly constrain interactions to the few-mode subspace.

The fact that the correctly truncated transverse electric field operator is expanded in terms of mode profiles  $\mathbf{f}'_\mu(\mathbf{x})$  which are a linear combination of the mode profiles (in the truncated basis) for the vector potential  $\mathbf{f}_\mu(\mathbf{x})$  is a fundamental feature of dissipation, and stands in contrast to the case of a normal mode expansion, presented in Appendix D. Nonetheless, under an assumption of well-separated discrete modes (e.g., high  $Q$ -factor resonators), the different modal expansions can be related to each other, allowing the transverse electric field to be expanded in the usual form, which we discuss in Appendix C.

The most general arbitrary-gauge Hamiltonian is then

$$\hat{\mathcal{H}}^g = \hat{\mathcal{V}}_g \hat{\mathcal{H}}_F \hat{\mathcal{V}}_g^\dagger + \hat{\mathcal{U}}_g \hat{\mathcal{H}}_0 \hat{\mathcal{U}}_g^\dagger - \int d^3x [\hat{\mathcal{E}}_F(\mathbf{x})]_\parallel \cdot \hat{\mathcal{P}}_\parallel(\mathbf{x}), \quad (69)$$

where we use a tilde to denote explicitly quantities with both photonic space truncation as well as material truncation. Equation (69) is, as expected, consistent with previous work on restoring gauge invariance under material and mode truncation [24,27,29]. Note that no loss of gauge invariance occurs if the longitudinal part of the medium-assisted electric field  $[\hat{\mathbf{E}}_F(\mathbf{x})]_{\parallel}$  is truncated (i.e., the part that belongs to the photonic subspace). If the transverse mode expansion is complete, in that it satisfies Eq. (64), the untruncated result is recovered. Note also that we can write the first two terms on the right-hand side of Eq. (69) as  $\hat{\mathcal{W}}_{g,C} \hat{\mathcal{H}}_F \hat{\mathcal{W}}_{g,C}^{\dagger} + \hat{\mathcal{W}}_{g,mp} \hat{\mathcal{H}}_0 \hat{\mathcal{W}}_{g,mp}^{\dagger}$ , which clearly indicates the special role that the Coulomb and multipolar gauges play in the theory of gauge invariance in truncated optical systems.

These expressions can be simplified in the Coulomb and multipolar gauges by noting that for the Coulomb gauge, we have  $\hat{\mathcal{V}}_C = 1$ , and in the multipolar gauge,

$$\int d^3x \mathbf{K}_{\perp}^{\text{mp}}(\mathbf{x}, \hat{\varepsilon}_{\alpha}) \cdot \mathbf{f}_{\mu}(\mathbf{x}) = -\hat{\varepsilon}_{\alpha} \cdot \int_0^1 ds \mathbf{f}_{\mu}(s\hat{\varepsilon}_{\alpha}). \quad (70)$$

A gauge transformation in the truncated photonic space is, similar to Eq. (57),

$$\hat{\mathcal{W}}_{g'g} = \exp \left[ -\frac{i}{\hbar} \int d^3x [\hat{\mathcal{P}}'_{\perp}(\mathbf{x}) - \hat{\mathcal{P}}_{\perp}(\mathbf{x})] \cdot \hat{\mathcal{A}}_{\perp}(\mathbf{x}) \right], \quad (71)$$

which again can be shown to be equivalent to a replacement of the truncated vector potential and transverse polarization. Of course, all the results in this section can easily be generalized to consider the case of no material truncation by taking  $\hat{\varepsilon}_{\alpha} \rightarrow \hat{\mathbf{r}}_{\alpha}$ .

The difference between Eq. (69), the correctly mode-truncated arbitrary-gauge Hamiltonian, and a naively truncated Hamiltonian  $\hat{P}_M \hat{\mathcal{H}}^g \hat{P}_M$  is twofold: First, the second term in Eq. (65), the interaction term between photonic and material subspaces, is expressed only in terms of the correctly truncated electric field if the Hamiltonian is properly (not directly) truncated, by instead truncating directly the vector potential as the fundamental field coordinate.

Second, the final term in Eq. (65) differs from the final term in Eq. (51) in that it contains explicit reference to the electromagnetic field via the mode functions  $\mathbf{f}_{\mu}$  [24]. In the case where truncation is applied after calculating the unitary transformation induced by  $\hat{\mathcal{V}}$ , the transverse delta function in the commutator (19a) (which requires a complete set of transverse modes) is used to remove any reference to the field modes. The additional integration present in the properly truncated case resolves issues with the  $\hat{\mathcal{P}}_{\perp}^2$  term when using the multipolar gauge, where  $\hat{\mathcal{P}}_{\perp}^{\text{mp}}$  is expressed in terms of a Dirac delta functions, which can cause problems due to the presence of a product of distributions [46].

It is instructive to compare our modal expansion and truncation in a fundamentally lossy and dispersive system, with the more commonly employed case in quantum optics of a *normal mode* expansion. To focus on the case of light-matter interactions in a medium, we can consider an inhomogeneous dielectric with a real and independent of frequency permittivity  $\epsilon(\mathbf{x})$ . In this case, the Helmholtz equation can be used to calculate the true normal modes of this system as the eigenfunctions of  $\nabla \times \nabla \times \mathbf{h}_{\mu} - \frac{\omega_{\mu}^2}{c^2} \epsilon(\mathbf{x}) \mathbf{h}_{\mu} = 0$ , with

appropriate closed or periodic boundary conditions. The modal eigenfunctions here are  $\mathbf{h}_{\mu}(\mathbf{x})$  with eigenvalue  $\omega_{\mu}$ , and they are generalized transverse in that they satisfy  $\nabla \cdot (\epsilon \mathbf{h}_{\mu}) = 0$ . In this case, an approximate (macroscopically averaged) Lagrangian can be used with no reference to the medium oscillator fields. In Appendix D we quantize this system using Dirac's constrained quantization procedure for the important case of the generalized Coulomb gauge, which satisfies  $\nabla \cdot (\epsilon \mathbf{A}) = 0$ , and the generalized multipolar gauge which we obtain by PZW transformation. The main result is that the generalized multipolar Hamiltonian  $\hat{\mathcal{H}}_{\text{gmp}}$  can be written as

$$\hat{\mathcal{H}}_{\text{gmp}} = \hat{\mathcal{H}}_0 + \sum_{\mu} \hbar \omega_{\mu} \hat{a}_{\mu}^{\dagger} \hat{a}_{\mu} - \int d^3x \hat{\mathcal{E}}_F(\mathbf{x}) \cdot \hat{\mathcal{Z}}_A(\mathbf{x}) + \sum_{\mu} \frac{[\int d^3x \hat{\mathcal{Z}}_A(\mathbf{x}) \cdot \mathbf{h}_{\mu}(\mathbf{x})]^2}{2\epsilon_0}, \quad (72)$$

where the part of the electric field operator which can be expressed in terms of bosonic normal mode operators is  $\hat{\mathcal{E}}_F(\mathbf{x}) = i \sum_{\mu} \sqrt{\frac{\hbar \omega_{\mu}}{2\epsilon_0}} \mathbf{h}_{\mu}(\mathbf{x}) \hat{a}_{\mu} + \text{H.c.}$ , and for each mode with profile  $\mathbf{h}_{\mu}$ , we can associate creation and annihilation operators which satisfy  $[\hat{a}_{\mu}, \hat{a}_{\nu}^{\dagger}] = \delta_{\mu\nu}$ . It is important to note that the full electric field operator also contains contributions from the material system operators, which must be considered when calculating physical observables; for details, see Appendix D.

One important difference between the generalized multipolar Hamiltonian (72) for a system supporting normal modes and the multipolar Hamiltonian (69), which can be formulated also for systems with lossy and nondiagonal mode expansions, is that the interaction term between photonic and material subspaces takes a form which includes only direct mode-polarization couplings in the former case, where loss can be neglected. As a result, a naive truncation of the generalized multipolar Hamiltonian (by applying  $\hat{P}_M$  operators directly to  $\hat{H}^{\text{mp}}$ ) differs from the correct result only in the polarization-squared term, which does not couple to the photonic subspace. In many cases (e.g., a fermionic TLS; see Sec. VIB), this term is irrelevant, or otherwise neglected. The ambiguity associated with mode truncation thus does not always play a significant role in light-matter interactions, at least with respect to the Hamiltonian. Note, however, that the generalized transverse eigenmodes are found with respect to the entire dielectric system and thus are generally delocalized and not appropriate for a discrete resonant mode truncation, despite this procedure being common in the literature (i.e., taking a “single-mode” limit).

In contrast, when considering a truncation of lossy modes, due to the cross-mode nature of the coupling terms in the second term of Eq. (65), truncation in the multipolar gauge will invariably break gauge invariance in a way which can have nonnegligible consequences. Our work generalizes previous results by Ref. [24] on restoring gauge invariance in truncated normal mode systems to consider these more general lossy mode expansions. One important example of this can be seen by considering open quantum systems.

To be concrete, consider the case of a cavity resonator that has discrete mode operators which couple to external fields and thus exhibits photon loss with some decay rate. Since

there is only one set of complete-space electromagnetic fields (i.e., of the universe), we propose that a rigorous model for the dynamics of the cavity field, where the other “reservoir” degrees of freedom are traced out (e.g., in the form of a master equation, using the well-known input-output formalism [52]), requires the use of the Coulomb gauge. This is because, in a generic gauge, the “reservoir” field which photons decay into is described by an entangled state of bosonic and fermionic subspaces, which prevents the use of standard open quantum system techniques to trace out the reservoir subsystem. However, the Coulomb gauge is *unique* in that it is the only gauge that satisfies  $\hat{\mathcal{V}} = 1$ . Thus, the unitary transformation which transforms the field Hamiltonian  $\hat{H}_F$ , and mixes up field and material degrees of freedom, becomes trivial in this gauge. The Coulomb gauge is thus the unique gauge (within the gauges realizable in the quantization function method) wherein the reservoir degrees of freedom are constrained entirely within a bosonic subspace, which is necessary to perform a Born-Markov approximation and derive a master equation.

It should of course be noted that one can perform, for example, a PZW-like transformation by using *only* the cavity mode degrees of freedom, to get something resembling the multipolar gauge for a truncated mode system. This is equivalent to defining a new class of gauge transformations which act only on the reduced system degrees of freedom. In this case, the reduced cavity system is by definition a mode-truncated system. Thus, the potential gauge ambiguity related to mode truncation, and the techniques described in this work to restore gauge invariance under mode truncation, are *intrinsic* to open quantum systems. The Hamiltonian in Eq. (69) (or an analogous construction) is generally required as a starting point to derive the correct model of open quantum system dynamics involving lossy cavities, if the system-reservoir coupling is to be derived rigorously (e.g., as in quantized QNM theory [41]).

Recently, it has been argued [62] that it is incorrect to assume that to model an open quantum system (e.g., a lossy cavity), the system operator that should couple to external “reservoir” modes is the transverse electric field operator, and implied that previous work by some of us [53] is incorrect as a result of this. This argument relies on the fact that, when matter degrees of freedom (e.g., “atoms” in the cavity) are localized far from the boundary of the resonator system, they should not play a role in the dissipation process of the cavity. Indeed, this is the case, and this feature is generally observed in rigorous models of loss from quantized cavity systems (e.g., [39,44,72,73]). It is a straightforward consequence of this that the electric field operator (or specifically, its modal excitation and de-excitation operators) is in fact the correct operator to use in describing coupling to a reservoir system, as it consists of only optical degrees of freedom. Other works rigorously treating dissipation in open cavities (albeit neglecting dispersion and absorption) have come to the same conclusion [44,73].

Where the potential confusion arises in Ref. [62], is that, when a complete set of modes is involved (no truncation), the arbitrary-gauge expression for the untruncated electric field operator is equivalent to that of its component which consists of only bosonic operators, for positions far away from the matter degrees of freedom (i.e., the cavity boundary). This led

the authors to suggest that the electric field evaluated away from the material degrees of freedom should be used to couple to the external reservoir, and that this field should be expanded using only boson operators. In a *correctly mode-truncated* picture, however, the electric field operator becomes  $\hat{\mathcal{E}}_{\perp}^g = \hat{\mathcal{V}}_g[\hat{\mathcal{E}}_F]_{\perp}\hat{\mathcal{V}}_g^{\dagger}$ , which cannot exclusively be described by bosonic degrees of freedom away from the material particle locations unless a complete set of modes is used to describe the cavity field. Using a single mode necessarily collapses the spatial degrees of freedom of the description of the electric field to a single coordinate, and thus the field at the boundary of the cavity and at the location of the material degrees of freedom are described by the same modal expansion operators. This was already pointed out in [59] (especially see Appendix A of same paper). Thus, by considering the correctly truncated form of the field operator, the intuition behind the argument of Ref. [62] does not hold, and consequently the gauge-invariant observables reported in [53] are correct within the assumptions of the model considered.

## V. TIME-DEPENDENT GAUGE TRANSFORMATIONS

In this section we generalize our approach to consider time-dependent gauge conditions and transformations, as well as how to construct gauge-invariant phenomenological time-dependent models of light-matter interaction.

We can extend the previously developed theory of arbitrary gauge quantization by allowing the transverse part of the quantization function to be an explicit function of time, such that (dropping the explicit  $g$  index for now)  $\mathbf{K} = \mathbf{K}(\mathbf{x}, \mathbf{x}', t)$ . In this case, one can follow the development in prior sections in the exact same way by taking the system to be quantized at a definite time  $t_0$ , and then determining the explicit time dependence of any observables when calculating expectation values [67,74].

Alternatively, we can account for the explicit time dependence of observables that arises from the time-dependent gauge condition by means of an additional time-dependent Hamiltonian term, generated by a canonical transformation prior to quantization, and then treating all operators as having no explicit time dependence. This is analogous to the approach in Ref. [59], where, *after* quantization, a unitary transformation with explicit time dependence is introduced to define a time-dependent gauge transformation. Here we generalize this approach by introducing it before quantization, and for arbitrary gauges.

Specifically, a modification to the Hamiltonian arises due to the canonical transformation that takes the coordinates  $(\mathbf{A}, \Pi_{\mathbf{A}})$  to  $(\mathbf{A}_{\perp}, \Pi)$ . With time-independent constraints, this transformation does not alter the Hamiltonian. With time-dependent constraints, however,  $\mathbf{P}(\mathbf{x})$  becomes explicitly time-dependent, and more care is needed to perform the canonical transformation to the unconstrained variables [74].

A canonical transformation from the set of constrained variables  $(\mathbf{A}, \Pi_{\mathbf{A}})$  to the unconstrained variables  $(\mathbf{A}_{\perp}, \Pi)$  can be implemented by writing the Lagrangian expressed in terms of the constrained variables as equal to a new Lagrangian expressed in terms of the unconstrained variables, up to a total time derivative which does not affect the extremization of the

action:

$$\begin{aligned} & \int d^3x \dot{\mathbf{A}} \cdot \boldsymbol{\Pi}_{\mathbf{A}} - H(\mathbf{A}, \boldsymbol{\Pi}_{\mathbf{A}}, t) \\ &= \int d^3x \dot{\mathbf{A}}_{\perp} \cdot \boldsymbol{\Pi} - H'(\mathbf{A}_{\perp}, \boldsymbol{\Pi}, t) + \frac{dG}{dt}. \end{aligned} \quad (73)$$

Here we have used a condensed notation by suppressing functional dependencies of the fields as well as the dependence of the Hamiltonians  $H$  and  $H'$  on the variables unaffected by the desired transformation. One way to implement this canonical transformation is by using a type-2 generating function [75]:

$$G = - \int d^3x \mathbf{A}_{\perp} \cdot \boldsymbol{\Pi} + \int d^3x G_2(\mathbf{A}, \boldsymbol{\Pi}, t). \quad (74)$$

For Hamilton's equations to be preserved in the unconstrained variables, one requires then  $\frac{\partial G_2}{\partial \mathbf{A}} = \boldsymbol{\Pi}_{\mathbf{A}}$ ,  $\frac{\partial G_2}{\partial \boldsymbol{\Pi}} = \mathbf{A}_{\perp}$ , and  $H' = H + \int d^3x \frac{\partial G_2}{\partial t}$ . Clearly, this can be accomplished by a form  $G_2 = \mathbf{A} \cdot \boldsymbol{\Pi}_{\mathbf{A}} + \boldsymbol{\Pi} \cdot \mathbf{A}_{\perp}$ , which we can express as a function of  $\mathbf{A}$  and  $\boldsymbol{\Pi}$  as

$$G_2(\mathbf{A}, \boldsymbol{\Pi}, t) = \mathbf{A} \cdot [\boldsymbol{\Pi} + \mathbf{P}(t)] + \boldsymbol{\Pi} \cdot \int d^3x' \delta^{\perp}(\mathbf{x} - \mathbf{x}') \cdot \mathbf{A}(\mathbf{x}'), \quad (75)$$

where we have emphasized that the explicit time dependence comes from (the transverse part of)  $\mathbf{P}$ . From this, it is easy to determine the relation

$$\begin{aligned} & H'(t) - H(t) \\ &= \int d^3x \frac{\partial G_2(\mathbf{A}, \boldsymbol{\Pi}, t)}{\partial t} \\ &= - \int d^3x \int d^3x' \mathbf{A}_{\perp}(\mathbf{x}, t) \cdot \left[ \frac{\partial}{\partial t} \mathbf{K}_{\perp}(\mathbf{x}, \mathbf{x}', t) \right] \rho_{\mathbf{A}}(\mathbf{x}', t). \end{aligned} \quad (76)$$

Upon quantization, this term becomes simply  $\int \hat{\mathbf{A}}_{\perp} \cdot \dot{\hat{\mathbf{P}}}_{\perp}$ , and so the only change in the case of time-dependent gauge function is

$$\hat{H} \rightarrow \hat{H}(t) = \hat{H} + \int d^3x \hat{\mathbf{A}}_{\perp}(\mathbf{x}) \cdot \dot{\hat{\mathbf{P}}}_{\perp}(\mathbf{x}, t). \quad (77)$$

This additional term added to the Hamiltonian is equivalent to the gauge transformation  $\hat{\phi} \rightarrow \hat{\phi} - \hat{\Lambda}$ , since this transforms  $[\hat{\mathbf{E}}_{\mathbf{F}}]_{\parallel} \rightarrow [\hat{\mathbf{E}}_{\mathbf{F}}]_{\parallel} + \hat{\Lambda}_{\parallel}$  (i.e., if  $\hat{\mathbf{A}}_{\parallel} = \nabla \hat{\Lambda}$ ), which when applied to the longitudinal interaction term in Eq. (69), and employing the quantization constraint relation in Eq. (9), generates the term in Eq. (77). Thus, a broad class of gauge transformations as defined in Eqs. (4a) and (4b) can be implemented within the arbitrary gauge quantization theory.

The notion of gauge *transformations* also must be modified to account for the explicit time dependence of the gauge transformation function,  $\hat{W}_{g'g}(t)$ . Here we shall consider the untruncated case for notational simplicity, although the procedure is identical with truncation of the material and/or photonic subspaces.

It is simple to show that the form of the Schrödinger evolution is conserved under a gauge transformation  $|\psi'_g(t)\rangle = \hat{W}_{g'g}(t) |\psi_g(t)\rangle$ , provided the Hamiltonian changes as

(restoring gauge indices)

$$\hat{H}^{g'}(t) = \hat{W}_{g'g}(t) \hat{H}^g(t) \hat{W}_{g'g}^{\dagger}(t) - i\hbar \hat{W}_{g'g}(t) \frac{\partial}{\partial t} \hat{W}_{g'g}^{\dagger}(t). \quad (78)$$

Using the definition of  $\hat{W}_{g'g}(t)$  [Eq. (30), but with  $\hat{\Lambda}_g(\mathbf{x}) \rightarrow \hat{\Lambda}_g(\mathbf{x}, t)$ ], one can show the second term in the above equation then takes the form

$$-i\hbar \hat{W}_{g'g}(t) \frac{\partial}{\partial t} \hat{W}_{g'g}^{\dagger}(t) = \int d^3x [\hat{\mathbf{P}}_{\perp}^{g'}(\mathbf{x}, t) - \hat{\mathbf{P}}_{\perp}^g(\mathbf{x}, t)] \cdot \hat{\mathbf{A}}_{\perp}(\mathbf{x}). \quad (79)$$

Reference [59] suggests that the Coulomb gauge is somehow more fundamental than the multipolar gauge, and must be used when time-dependent interactions are to be considered. Here, we take the view that gauge symmetry is a fundamental property of the QED Lagrangian, and thus any gauge should give consistent results, provided any approximations to the theory are implemented consistently across gauges.

When describing time-dependent interactions and/or gauge conditions (as in Refs. [59,60]), however, the Coulomb gauge is potentially *unique* in that it is defined, conveniently, by  $\mathbf{K}_{\perp}^{\mathbf{C}}(\mathbf{x}, \mathbf{x}', t) = 0$  for all  $t$ , and thus is a time-invariant gauge. In contrast, when the light-matter interaction strength is given an explicit time dependence, the multipolar gauge constraint condition itself should also be given this explicit time dependence to get something similar to the usual form of the multipolar Hamiltonian. Thus, the additional term [Eq. (77)] arises, which does not exist in the time-independent multipolar gauge.

It should also be noted that an explicitly time-dependent Hamiltonian is always an approximation, since the energy nonconserving nature of a time-dependent Hamiltonian means some external system has dynamics which are not explicitly modeled. Thus, choosing to impose time dependence in a specific gauge can potentially lead to different results than if it is imposed in other gauges, and analysis of the physical origin of the time dependence may need to be taken to resolve this potential ambiguity [60].

For the case of time-dependent coupling between the transverse field and matter, we can introduce this in a way which is unambiguous and equivalent in both the Coulomb and multipolar gauges as follows: in the Coulomb gauge, the minimal coupling is introduced via the transverse part of  $\hat{\mathbf{Z}}_{\mathbf{A}}$  [see Eq. (43)]. Thus, we can, for example, modulate the transverse field's interaction strength with matter by phenomenologically modulating this parameter by a time-dependent factor  $\mu(t)$  such that  $\hat{\mathbf{Z}}_{\mathbf{A}}(\mathbf{x}) \rightarrow \mu(t) \hat{\mathbf{Z}}_{\mathbf{A}}(\mathbf{x})$ . In the multipolar gauge, the coupling is entirely mediated by the transverse polarization  $\hat{\mathbf{P}}_{\perp}^{\text{mp}}(\mathbf{x})$ . Thus, we can make the equivalent approximation by taking  $\hat{\mathbf{P}}_{\perp}^{\text{mp}}(\mathbf{x}) \rightarrow \mu(t) \hat{\mathbf{P}}_{\perp}^{\text{mp}}(\mathbf{x})$ .

Now, in the Coulomb gauge,  $\hat{\mathbf{Z}}_{\mathbf{A}}(\mathbf{x})$  is simply a function which generates (neglecting the magnetic interaction terms) the minimal coupling replacements  $\hat{\mathbf{p}}_{\alpha} \rightarrow \hat{\mathbf{p}}_{\alpha} - q_{\alpha} \hat{\mathbf{A}}_{\perp}(\hat{\mathbf{r}}_{\alpha})$ . In contrast, in the multipolar gauge,  $\hat{\mathbf{P}}_{\perp}^{\text{mp}}(\mathbf{x})$  is a parameter which appears in the Lagrangian. As such, by making this quantity time-dependent, we must add to the Hamiltonian the additional term given by Eq. (77) and understand the gauge condition as depending explicitly on time. Additionally, one must still make the replacement  $[\hat{\mathbf{Z}}_{\mathbf{A}}(\mathbf{x})]_{\perp} \rightarrow \mu(t) [\hat{\mathbf{Z}}_{\mathbf{A}}(\mathbf{x})]_{\perp}$

to ensure  $\hat{U}_{\text{mp}} = 1$  even under the new time-dependent gauge condition. In summary, we can implement time-dependent phenomenological interactions between the transverse field and the active material particles in any gauge by letting

$$[\hat{\mathbf{Z}}_{\text{A}}(\mathbf{x})]_{\perp} \rightarrow \mu(t)[\hat{\mathbf{Z}}_{\text{A}}(\mathbf{x})]_{\perp}, \quad (80a)$$

$$\mathbf{K}_{\perp}^g(\mathbf{x}, \mathbf{x}') \rightarrow \mathbf{K}_{\perp}^g(\mathbf{x}, \mathbf{x}', t) = \mu(t)\mathbf{K}_{\perp}^g(\mathbf{x}, \mathbf{x}'), \quad (80b)$$

$$\hat{H} \rightarrow \hat{H} + \dot{\mu}(t) \int d^3x \hat{\mathbf{A}}_{\perp}(\mathbf{x}) \cdot \hat{\mathbf{P}}_{\perp}^g(\mathbf{x}). \quad (80c)$$

To transform between gauges, we can perform the time-dependent gauge transformation via  $\hat{W}_{\text{C,mp}}(t)$  [see Eq. (30)], following the rule in Eq. (78)] to transform from one gauge to another, and the additional term is correctly accounted for by this transformation. Consequently, one can introduce time-dependent interactions in systems with ultrastrong coupling in either gauge without introducing any ambiguity or violating gauge invariance. Using our formalism of the time-dependent quantization function  $\mathbf{K}(\mathbf{x}, \mathbf{x}', t)$ , this is done without treating any gauge as more fundamental than another, although as noted in Ref. [59], this is most easily done in the Coulomb gauge, where the coupling strength can straightforwardly be made time-dependent.

In contrast to Ref. [60], which claims that introducing time-dependent light-matter interaction strengths necessarily breaks the gauge invariance of the fundamental light-matter Lagrangian, we note that this is circumvented by applying the time-dependent modulation of the interaction strength only to the transverse part of the interaction, since the transverse vector potential is gauge-invariant. This in fact produces equivalent results to the replacements made at the level of the Coulomb and multipolar gauge Hamiltonians described above, and allows one to introduce gauge-invariant time-dependent couplings in a completely unambiguous way.

Thus ultrastrong time-dependent light-matter interactions are *not* gauge relative, as fully expected, if using a correct and unambiguous theory. Furthermore, it should also be noted that modulating the longitudinal (quasistatic) interaction is likely to lead to undesired and unphysical predictions beyond just the breaking of gauge invariance, as this would involve the modulation of the Coulomb forces which bind together the constituent atoms of the matter degrees of freedom. Clearly the desired phenomenological model in many cases should be that of a time-dependent transverse coupling only.

## VI. THE DIPOLE APPROXIMATION AND TWO-LEVEL SYSTEMS

In this section we apply the results of Sec. III to some commonly used models and approximations in quantum optics. Specifically, we discuss the *dipole approximation* (or long-wavelength approximation) in Sec. VIA, TLSs in Sec. VIB, and how to go beyond the dipole approximation for effective single-particle models in Sec. VIC. We focus on the case of the dispersive and absorbing dielectric, but analogous results for the case of a real dielectric, where the fields are expanded in terms of the generalized transverse normal modes of the system (see Appendix D) can easily be derived by applying the appropriate approximations to the generalized Coulomb and multipolar gauges in Eqs. (D39) and (D40).

### A. Dipole approximation

In the Coulomb and multipolar gauges, the Hamiltonian is in a particularly convenient form to expand the field potential functions around  $\mathbf{x} = \mathbf{r}_{\text{A}} = 0$ , the center of the, e.g., molecular charge distribution. In particular, taking this expansion to first order in  $\hat{\varepsilon}_{\alpha}$  results in the *dipole approximation*. For example, under the dipole approximation, the operator  $\hat{U}_g$  becomes (with no mode truncation) [27,29,59]

$$\hat{U}_g = \exp\left(\frac{i}{\hbar} \hat{\mathbf{d}} \cdot \hat{\mathbf{A}}_0^g\right), \quad (81)$$

where we have defined the *dipole operator*  $\hat{\mathbf{d}} = \sum_{\alpha} q_{\alpha} \hat{\varepsilon}_{\alpha}$ . Similarly,  $\hat{V}_{\text{mp}}$  becomes  $\hat{V}_{\text{mp}} = \exp(-\frac{i}{\hbar} \hat{\mathbf{d}} \cdot \hat{\mathbf{A}}_{\perp,0})$ . Note one could also define a dipole operator for each particle, which is naturally more suited to multiparticle models such as the Dicke and Hopfield models [76]. In this section we shall use fields with the 0 subscript to correspond to the evaluation at the origin (or, more generally at the center of a charge distribution  $\mathbf{x} = \mathbf{r}_{\text{A}}$ ); for example, here  $\hat{\mathbf{A}}_0 \equiv \hat{\mathbf{A}}(\mathbf{0})$ .

The Hamiltonian in the Coulomb gauge, after applying the dipole approximation to Eq. (52) (with no mode truncation), is

$$\begin{aligned} \hat{\mathcal{H}}^{\text{C}} &= \exp\left(\frac{i}{\hbar} \hat{\mathbf{d}} \cdot \hat{\mathbf{A}}_{\perp,0}\right) \hat{\mathcal{H}}_0 \exp\left(-\frac{i}{\hbar} \hat{\mathbf{d}} \cdot \hat{\mathbf{A}}_{\perp,0}\right) \\ &+ \hat{H}_{\text{F}} - \hat{\mathbf{d}} \cdot [\hat{\mathbf{E}}_{\text{F}}]_{\parallel,0}, \end{aligned} \quad (82)$$

and in the multipolar (or dipole) gauge

$$\hat{\mathcal{H}}^{\text{mp}} = \hat{H}_{\text{F}} + \hat{\mathcal{H}}_0 - \hat{\mathbf{d}} \cdot \hat{\mathbf{E}}_{\text{F},0} + \sum_{\alpha, \alpha'} \frac{q_{\alpha} q_{\alpha'}}{2\epsilon_0} \hat{\varepsilon}_{\alpha} \cdot \delta^{\perp}(\mathbf{0}) \cdot \hat{\varepsilon}_{\alpha'}. \quad (83)$$

The last term in Eq. (83) is not well-defined, which is a well-known problem [24,46]. However, under mode truncation, the divergence becomes finite:

$$\begin{aligned} \hat{\mathcal{H}}^{\text{mp}} &= \hat{\mathcal{H}}_{\text{F}} + \hat{\mathcal{H}}_0 + \left( i \sum_{\mu\nu} \hbar \chi_{\mu\nu}^* \hat{a}_{\mu} \hat{\xi}_{\nu}^{g\dagger} + \text{H.c.} \right) \\ &+ \sum_{\mu\nu} \hbar \chi_{\mu\nu} \hat{\xi}_{\mu}^{g\dagger} \hat{\xi}_{\nu}^g - \hat{\mathbf{d}} \cdot [\hat{\mathbf{E}}_{\text{F}}]_{\parallel,0}, \end{aligned} \quad (84)$$

where, within the dipole approximation,

$$\hat{\xi}_{\mu}^g = -\frac{\hat{\mathbf{d}} \cdot \mathbf{f}_{\mu}^*(\mathbf{0})}{\sqrt{2\epsilon_0 \hbar \chi_{\mu\mu}}}, \quad (85)$$

and the mode-truncated result for the Coulomb gauge simply replaces  $\hat{\mathbf{A}}_{\perp,0}$  with  $\hat{\mathcal{A}}_{\perp,0}$ . Note that, as mentioned in Sec. IV C and further justified in Sec. VII, the third term on the right-hand side of Eq. (84) can also be written as  $-\hat{\mathbf{d}} \cdot [\hat{\mathcal{E}}_{\text{F}}]_{\perp,0}$ , where  $[\hat{\mathcal{E}}_{\text{F}}]_{\perp}(\mathbf{x})$  is the *correctly truncated* transverse and bosonic part of the electric field operator defined in Eq. (68).

### B. Two-level systems

If we restrict ourselves to only two quantized material states  $|e\rangle$  and  $|g\rangle$  (for the active media), with energies  $+\hbar\omega_0/2$  and  $-\hbar\omega_0/2$ , respectively, then we can take advantage of the SU(2) Pauli matrix algebra, and write  $\hat{\mathcal{H}}_0 = \hbar\omega_0 \hat{\sigma}_z/2$ , where

$\hat{\sigma}_z = |e\rangle\langle e| - |g\rangle\langle g|$ . Furthermore, the only information about the TLS which is required, in addition to its energy level separation, is its dipole moment. We shall assume the two states to have parity symmetry, and take the off-diagonal matrix element to be real:

$$\langle e|\hat{\mathbf{d}}|e\rangle = \langle g|\hat{\mathbf{d}}|g\rangle = 0, \quad (86a)$$

$$\langle e|\hat{\mathbf{d}}|g\rangle = \langle g|\hat{\mathbf{d}}|e\rangle = \mathbf{d}, \quad (86b)$$

such that  $\hat{\mathcal{D}} = \mathbf{d}\hat{\sigma}_x$ , where  $\hat{\sigma}_x = |e\rangle\langle g| + |g\rangle\langle e|$ . We can then calculate the matrix exponentials in Eq. (82) to obtain the Coulomb gauge Hamiltonian under the dipole approximation for a TLS:

$$\begin{aligned} \hat{\mathcal{H}}_C = & \hat{H}_F - \mathbf{d} \cdot [\hat{\mathbf{E}}_F]_{\parallel,0} \hat{\sigma}_x \\ & + \frac{\hbar\omega_0}{2} \left[ \cos\left(\frac{2\mathbf{d} \cdot \hat{\mathbf{A}}_{\perp,0}}{\hbar}\right) \hat{\sigma}_z + \sin\left(\frac{2\mathbf{d} \cdot \hat{\mathbf{A}}_{\perp,0}}{\hbar}\right) \hat{\sigma}_y \right], \end{aligned} \quad (87)$$

where  $\hat{\sigma}_y = i|g\rangle\langle e| - i|e\rangle\langle g|$ . The corresponding multipolar gauge Hamiltonian is also easily obtained:

$$\hat{\mathcal{H}}_{\text{mp}} = \hat{H}_F + \frac{\hbar\omega_0}{2} \hat{\sigma}_z - \mathbf{d} \cdot \hat{\mathbf{E}}_{F,0} \hat{\sigma}_x. \quad (88)$$

In Eq. (88) we have been able to drop the problematic divergent term  $\mathbf{d} \cdot \delta^{-1}(\mathbf{0}) \cdot \mathbf{d}/(2\epsilon_0)$ , which although not well-defined, is a  $c$ -number which does not contribute to the Hamiltonian dynamics in the two-level subspace. The mode-truncated Hamiltonian in the TLS approximation is easily obtained by expressing Eq. (87) in terms of the truncated field  $\hat{\mathbf{A}}_{\perp,0}$  for the Coulomb gauge, and letting  $\hat{\mathcal{D}} \rightarrow \mathbf{d}\hat{\sigma}_x$  in Eq. (84). For example, in the single-mode limit where  $\chi_{\mu\nu} = \chi$ , we have, in alignment with previous works [27,53,59],

$$\begin{aligned} \hat{\mathcal{H}}_{(1)}^C = & \hbar\chi \hat{a}^\dagger \hat{a} - \mathbf{d} \cdot [\hat{\mathbf{E}}_F]_{\parallel,0} \hat{\sigma}_x + \frac{\hbar\omega_0}{2} \\ & \times \{ \cos[2(\eta\hat{a}^\dagger + \eta^*\hat{a})] \hat{\sigma}_z + \sin[2(\eta\hat{a}^\dagger + \eta^*\hat{a})] \hat{\sigma}_y \} \end{aligned} \quad (89)$$

and

$$\begin{aligned} \hat{\mathcal{H}}_{(1)}^{\text{mp}} = & \hbar\chi \hat{a}^\dagger \hat{a} - \mathbf{d} \cdot [\hat{\mathbf{E}}_F]_{\parallel,0} \hat{\sigma}_x \\ & + i\hbar\chi(\eta\hat{a}^\dagger - \eta^*\hat{a})\hat{\sigma}_x + \frac{\hbar\omega_0}{2} \hat{\sigma}_z, \end{aligned} \quad (90)$$

where  $\eta = \mathbf{d} \cdot \mathbf{f}^*(\mathbf{0})/\sqrt{2\epsilon_0\hbar\chi}$ , and we have dropped a term proportional to the identity.

### C. Beyond the dipole approximation for a two-level system

Here we discuss an effective single-particle model for a TLS. The results in this section can be generalized to higher-level systems as well.

To derive general results beyond the dipole approximation, one can assume in some instances an effective single-particle model for the polarization consisting of a single charge  $q$  with position operator  $\hat{\mathbf{x}}$ , and a charge  $-q$  which remains in a position eigenstate of the Hamiltonian with eigenvalue  $\mathbf{r}_A = 0$ . This is just one toy model (reminiscent of the hydrogen atom), and other effective single-particle models exist; for example, we could also have a charge  $+q$  with position operator  $\hat{\mathbf{x}}$  and a

charge  $-q$  with position operator  $-\hat{\mathbf{x}}$ . Sticking with the former case, we have

$$\hat{\mathcal{P}}^g(\mathbf{x}) = -q\mathbf{K}^g(\mathbf{x}, \hat{\mathbf{x}}), \quad (91)$$

where  $\hat{\mathbf{x}} = \mathbf{d}\hat{\sigma}_x/q$ . We can derive simpler expressions by noting that, since  $\hat{\sigma}_x$  is proportional to the position operator, one can immediately evaluate any function of operators which depends only on the position operator by using the eigenbasis of  $\hat{\sigma}_x$ . By construction (to preserve gauge invariance upon truncation), we have formulated our theory such that the entire Hamiltonian can be simplified in this manner. For example, for a function  $\hat{f}(\hat{\mathbf{x}})$ , one can write

$$\hat{f}(\hat{\mathbf{x}}) = \frac{f(\mathbf{r}_{\text{dip}}) + f(-\mathbf{r}_{\text{dip}})}{2} + \frac{f(\mathbf{r}_{\text{dip}}) - f(-\mathbf{r}_{\text{dip}})}{2} \hat{\sigma}_x, \quad (92)$$

where  $\mathbf{r}_{\text{dip}} = \mathbf{d}/q$  is an effective position coordinate. Such a representation could be of course generalized to include cases where the position operator is a generic Hermitian matrix in the two-level subspace. We find, not yet considering mode truncation,

$$\begin{aligned} \hat{\mathcal{H}}^g = & \hat{H}_F + \frac{\hbar\omega_0}{2} [\cos(\hat{\Phi}_A^g) \hat{\sigma}_z + \sin(\hat{\Phi}_A^g) \hat{\sigma}_y] \\ & + \frac{q^2}{4\epsilon_0} \int d^3x \{ [\mathbf{K}_\perp^g(\mathbf{x}, \mathbf{r}_{\text{dip}})]^2 - [\mathbf{K}_\perp^g(\mathbf{x}, -\mathbf{r}_{\text{dip}})]^2 \} \hat{\sigma}_x \\ & + \frac{q}{2} \int d^3x \hat{\mathbf{E}}_F(\mathbf{x}) \cdot [\mathbf{K}^g(\mathbf{x}, \mathbf{r}_{\text{dip}}) + \mathbf{K}^g(\mathbf{x}, -\mathbf{r}_{\text{dip}})] \\ & + \frac{q}{2} \int d^3x \hat{\mathbf{E}}_F(\mathbf{x}) \cdot [\mathbf{K}^g(\mathbf{x}, \mathbf{r}_{\text{dip}}) - \mathbf{K}^g(\mathbf{x}, -\mathbf{r}_{\text{dip}})] \hat{\sigma}_x, \end{aligned} \quad (93)$$

where  $\hat{\Phi}_A^g = \mathbf{d} \cdot \int_{-1}^1 ds \hat{\mathbf{A}}^g(s\mathbf{r}_{\text{dip}})/\hbar$ , and we have dropped the  $c$ -number term,

$$\frac{q^2}{4\epsilon_0} \int d^3x \{ [\mathbf{K}_\perp^g(\mathbf{x}, \mathbf{r}_{\text{dip}})]^2 + [\mathbf{K}_\perp^g(\mathbf{x}, -\mathbf{r}_{\text{dip}})]^2 \}, \quad (94)$$

which does not contribute dynamically. In the case of mode truncation, the Hamiltonian  $\hat{\mathcal{H}}^g$  is similar to Eq. (93), but with replacements  $\hat{H}_F \rightarrow \hat{\mathcal{H}}_F$ ,  $\hat{\Phi}_A^g \rightarrow \hat{\Phi}_{\mathcal{A}}^g$ , and  $\hat{\mathbf{E}}_F \rightarrow \hat{\mathcal{E}}_F = [\hat{\mathcal{E}}_F]_\perp + [\hat{\mathbf{E}}_F]_\parallel$ , where  $[\hat{\mathcal{E}}_F]_\perp$  is the bosonic portion of the correctly truncated transverse electric field operator, defined in Eq. (68). Additionally, the second line instead becomes

$$\begin{aligned} \frac{q^2}{4\epsilon_0} \sum_\mu \left\{ \left[ \int d^3x \mathbf{K}_\perp^g(\mathbf{x}, \mathbf{r}_{\text{dip}}) \cdot \mathbf{f}_\mu(\mathbf{x}) \right]^2 \right. \\ \left. - \left[ \int d^3x \mathbf{K}_\perp^g(\mathbf{x}, -\mathbf{r}_{\text{dip}}) \cdot \mathbf{f}_\mu(\mathbf{x}) \right]^2 \right\} \hat{\sigma}_x, \end{aligned} \quad (95)$$

and the  $c$ -number term that is dropped is modified accordingly to smooth out the divergence.

From this result, it is straightforward to obtain the Coulomb gauge Hamiltonian beyond the dipole approximation:

$$\begin{aligned} \hat{H}^C &= \hat{H}_F + \frac{\hbar\omega_0}{2} [\cos(\hat{\Phi}_A^C)\hat{\sigma}_z + \sin(\hat{\Phi}_A^C)\hat{\sigma}_y] \\ &\quad - \frac{\mathbf{d}}{2} \cdot \int_0^1 ds \{ [\hat{\mathbf{E}}_F(s\mathbf{r}_{\text{dip}})]_{\parallel} - [\hat{\mathbf{E}}_F(-s\mathbf{r}_{\text{dip}})]_{\parallel} \} \\ &\quad - \frac{\mathbf{d}}{2} \cdot \int_{-1}^1 ds [\hat{\mathbf{E}}_F(s\mathbf{r}_{\text{dip}})]_{\parallel} \hat{\sigma}_x. \end{aligned} \quad (96)$$

Similarly, for the multipolar gauge,

$$\begin{aligned} \hat{H}_{\text{mp}} &= \hat{H}_F + \frac{\hbar\omega_0}{2} \hat{\sigma}_z \\ &\quad - \frac{\mathbf{d}}{2} \cdot \int_0^1 ds [\hat{\mathbf{E}}_F(s\mathbf{r}_{\text{dip}}) - \hat{\mathbf{E}}_F(-s\mathbf{r}_{\text{dip}})] \\ &\quad - \frac{\mathbf{d}}{2} \cdot \int_{-1}^1 ds \hat{\mathbf{E}}_F(s\mathbf{r}_{\text{dip}}) \hat{\sigma}_x. \end{aligned} \quad (97)$$

Note that the second line of Eqs. (96) and (97) vanish for the common situation in which the electric field is an even function of position (e.g., in cavity QED where a dipole is placed at a modal antinode). With mode truncation, Eqs. (96) and (97) should be expressed in terms of the truncated field variables, with the additional term to the multipolar gauge Hamiltonian:

$$\frac{1}{4\epsilon_0} \sum_{\mu} \left\{ \left[ \mathbf{d} \cdot \int_0^1 ds \mathbf{f}_{\mu}(s\mathbf{r}_{\text{dip}}) \right]^2 - \left[ \mathbf{d} \cdot \int_{-1}^0 ds \mathbf{f}_{\mu}(s\mathbf{r}_{\text{dip}}) \right]^2 \right\} \hat{\sigma}_x. \quad (98)$$

Considering a one-dimensional system and neglecting the longitudinal field terms, the Coulomb gauge result shown by Eq. (96) was also found in Ref. [31], where it was noted that due to the spatial integral over the transverse vector potential, going beyond the dipole approximation introduces a natural cutoff for high-frequency interactions, as  $\hat{\Phi}_A^C$  vanishes for mode wavelengths much shorter than  $|\mathbf{r}_{\text{dip}}|$ . Here we have extended these results to consider an arbitrary gauge Hamiltonian for the general three-dimensional case, and including the longitudinal terms.

## VII. RESOLUTION OF GAUGE AMBIGUITY ASSOCIATED WITH PHOTON DETECTION AND FIELD OBSERVABLES

A potential ambiguity that can arise when determining *observables* of the electromagnetic field is the gauge-dependent nature of the field operators. As an example, consider the transverse electric field operator in the Coulomb gauge  $\hat{\mathbf{E}}_{\perp}^C = [\hat{\mathbf{E}}_F]_{\perp}$ , and the multipolar gauge,  $\hat{\mathbf{E}}_{\perp}^{\text{mp}} = [\hat{\mathbf{E}}_F]_{\perp} - \frac{1}{\epsilon_0} \hat{\mathbf{P}}_{\perp}^{\text{mp}}$ . In this section we focus on gauge ambiguities associated with (electromagnetic) mode truncation, so the polarization can be truncated or untruncated.

Suppose we want to calculate an observable which is a function of the electric field,  $\langle \hat{O}(\hat{\mathbf{E}}^g) \rangle_g$ , for gauges  $g = C, \text{mp}$ , and the subscript on the expectation value indicates it is to be calculated with respect to a state vector (or density operator) in the gauge  $g$ . Without any truncation, we have, by construction of the manifestly gauge invariant quantum theory presented in previous sections,  $\langle \hat{O}(\hat{\mathbf{E}}^C) \rangle_C = \langle \hat{O}(\hat{\mathbf{E}}^{\text{mp}}) \rangle_{\text{mp}}$ . Since  $\hat{\mathbf{P}}$

is nonzero only in the vicinity of the matter charged particles, one should be able to calculate any observable evaluated at locations far from the position of the free matter particles as  $\langle \hat{O}(\hat{\mathbf{E}}^C) \rangle_C = \langle \hat{O}(\hat{\mathbf{E}}^{\text{mp}}) \rangle_{\text{mp}} = \langle \hat{O}([\hat{\mathbf{E}}_F]_{\perp}) \rangle_C = \langle \hat{O}([\hat{\mathbf{E}}_F]_{\perp}) \rangle_{\text{mp}}$ . However, as the multipolar and Coulomb gauges have different Hamiltonians,  $\langle \hat{O}([\hat{\mathbf{E}}_F]_{\perp}) \rangle_C$  and  $\langle \hat{O}([\hat{\mathbf{E}}_F]_{\perp}) \rangle_{\text{mp}}$  will generally be different when mode truncation is considered.

To illustrate this issue, consider the following expansion of the bosonic part of the electric field:

$$[\hat{\mathbf{E}}_F(\mathbf{x})]_{\perp} = i \sum_{\mu} \sqrt{\frac{\hbar\chi_{\mu\mu}}{2\epsilon_0}} \mathbf{f}_{\mu}^E(\mathbf{x}) \hat{a}_{\mu} + \text{H.c.}, \quad (99)$$

where  $\mathbf{f}_{\mu}^E$  are proposed modal functions for the electric field expansion defined through naive projection, given in Appendix C. We will ultimately show that this is generally not the correct form of the electric field operator mode expansion when mode truncation is to be performed.

Suppose Eq. (99) is expressed in the Coulomb gauge, such that  $[\hat{\mathbf{E}}_F]_{\perp} = \hat{\mathbf{E}}_{\perp}$ . Then, to obtain the multipolar gauge expression for the transverse electric field, we apply the PZW transformation (using the dipole approximation for simplicity):

$$\hat{W}_{\text{mp},C} \hat{a}_{\mu} \hat{W}_{\text{mp},C}^{\dagger} = \hat{a}_{\mu} + i \frac{\hat{\mathbf{d}} \cdot \mathbf{f}_{\mu}^*(\mathbf{0})}{\sqrt{2\hbar\epsilon_0\chi_{\mu\mu}}}, \quad (100)$$

and so

$$\begin{aligned} &\hat{W}_{\text{mp},C} [\hat{\mathbf{E}}_F(\mathbf{x})]_{\perp} \hat{W}_{\text{mp},C}^{\dagger} \\ &= [\hat{\mathbf{E}}_F(\mathbf{x})]_{\perp} - \frac{\hat{\mathbf{d}}}{\epsilon_0} \cdot \left[ \frac{1}{2} \sum_{\mu} \mathbf{f}_{\mu}^E(\mathbf{x}) \mathbf{f}_{\mu}^*(\mathbf{0}) + \text{c.c.} \right]. \end{aligned} \quad (101)$$

The quantity in square brackets is equal to  $\delta^{\perp}(\mathbf{x})$  if the mode expansion is complete (see Appendix C), but for a finite mode expansion, it becomes nonzero even away from the dipole location at  $\mathbf{x} = \mathbf{0}$ . Thus, upon mode truncation we have a potential ambiguity; should we truncate with respect to the photonic Hilbert space, using  $\hat{P}_M$ , or with respect to the sum over mode index  $\mu$ ?

In Ref. [59] this potential ambiguity was identified, with the proposed solution to truncate with respect to the mode index. However, no definitive argument was given for why this should be the case, beyond the fact that it is unitarily equivalent to the seemingly less ambiguous situation in the Coulomb gauge. Here we show that this can be justified by a proper restoration of the gauge invariance lost under naive mode truncation.

To resolve this ambiguity, we take an approach similar in spirit to that of Ref. [59], by explicitly modeling the detector degree of freedom, and using the mode truncation procedure outlined in this work to derive a gauge-invariant model of the detector constituent particles.

To do so, we can apply the theory in previously developed sections (and the Appendixes) by generalizing the quantum charge and current densities to become  $\hat{\rho}_A(\mathbf{x}) \rightarrow \hat{\rho}_A(\mathbf{x}) + \hat{\rho}_d(\mathbf{x})$ , and  $\hat{\mathbf{J}}_A(\mathbf{x}) \rightarrow \hat{\mathbf{J}}_A(\mathbf{x}) + \hat{\mathbf{J}}_d(\mathbf{x})$ ,

where

$$\hat{\rho}_d(\mathbf{x}) = \sum_{\alpha_d} q_{\alpha_d} \delta(\mathbf{x} - \hat{\mathbf{r}}_{\alpha_d}), \quad (102)$$

$$\hat{\mathbf{J}}_d(\mathbf{x}) = \sum_{\alpha_d} q_{\alpha_d} \hat{\mathbf{r}}_{\alpha_d} \delta(\mathbf{x} - \hat{\mathbf{r}}_{\alpha_d}), \quad (103)$$

corresponding to “detector” particles with position operators  $\hat{\mathbf{r}}_{\alpha_d}$  and charge  $q_{\alpha_d}$  where again we assume  $\sum_{\alpha_d} q_{\alpha_d} = 0$ . This change induces  $\hat{\mathbf{P}}(\mathbf{x}) \rightarrow \hat{\mathbf{P}}(\mathbf{x}) + \hat{\mathbf{P}}_d(\mathbf{x})$ , to reflect the additional detector degrees of freedom in the polarization. We assume these particles to be localized around a position  $\mathbf{r}_d$ , and as such, we take the following form for the multipolar gauge polarization [as well as the corresponding change in  $\mathbf{Z}_A(\mathbf{x})$ ]:

$$\hat{\mathbf{P}}_d^{\text{mp}}(\mathbf{x}) = \sum_{\alpha_d} q_{\alpha_d} (\hat{\mathbf{r}}_{\alpha_d} - \mathbf{r}_d) \int_0^1 ds \delta[\mathbf{x} - \mathbf{r}_d - s(\hat{\mathbf{r}}_{\alpha_d} - \mathbf{r}_d)]. \quad (104)$$

Obviously this is an oversimplified model of photodetection [77], but it is sufficient to capture the essential features with respect to gauge invariance and mode truncation. For the sake of simplicity (and because we assume weak coupling of the detector to the field), we will assume a truncation of the total material subspace such that the detector degree of freedom can be reduced to a TLS with energy separation  $\hbar\omega_d$ . Similarly, we assume the dipole approximation at the location of the detector  $\mathbf{r}_d$ .

The Hamiltonian for this system under mode truncation is that of Eq. (69), generalized to incorporate the additional detector. Expanding to first order in the detector dipole moment operator  $\hat{\mathbf{d}}_d = \sum_{\alpha_d} q_{\alpha_d} (\hat{\mathbf{r}}_{\alpha_d} - \mathbf{r}_d)$ , we obtain the Coulomb and multipolar gauge Hamiltonians:

$$\hat{\mathcal{H}}_d^C = \hat{\mathcal{H}}^C + \frac{\hbar\omega_d}{2} \hat{\sigma}_z^d + \omega_d \mathbf{d}_d \cdot \hat{\mathcal{A}}_{\perp}(\mathbf{r}_d) \hat{\sigma}_y^d, \quad (105)$$

$$\hat{\mathcal{H}}_d^{\text{mp}} = \hat{\mathcal{H}}^{\text{mp}} + \frac{\hbar\omega_d}{2} \hat{\sigma}_z^d - \left( i \sum_{\mu\nu} \hbar\chi_{\mu\nu}^* \eta_{\nu}^{d*} \hat{a}'_{\mu} + \text{H.c.} \right) \hat{\sigma}_x^d, \quad (106)$$

where we have used  $\hat{\mathbf{d}}_d = \mathbf{d}_d \hat{\sigma}_x^d$ , and we have defined  $\eta_{\mu}^d = \mathbf{d}_d \cdot \mathbf{f}_{\mu}^*(\mathbf{r}_d) / (\sqrt{2\epsilon_0 \hbar \chi_{\mu\mu}})$  and  $\hat{a}'_{\mu} = \hat{a}_{\mu} + i\eta_{\mu} \hat{\sigma}_x$ , where  $\eta_{\mu} = \mathbf{d} \cdot \mathbf{f}_{\mu}^*(\mathbf{0}) / (\sqrt{2\epsilon_0 \hbar \chi_{\mu\mu}})$ . Suppose, now, that the detector TLS frequency  $\omega_d$  is resonant with a transition of the Hamiltonian (neglecting the detector part) between eigenstates  $|j\rangle$  and  $|i\rangle$  with frequency  $\omega_j - \omega_i = \omega_d$  [59]. By using perturbation theory (Fermi’s golden rule), the photodetection rate should be proportional to  $R_C$  and  $R_{\text{mp}}$  in the Coulomb and multipolar gauges, respectively, where

$$R_C = |\langle i_C | \omega_d \mathbf{d}_d \cdot \hat{\mathcal{A}}_{\perp}(\mathbf{r}_d) | j_C \rangle|^2, \quad (107)$$

$$R_{\text{mp}} = |\langle i_{\text{mp}} | i \sum_{\mu\nu} \hbar\chi_{\mu\nu}^* \eta_{\nu}^{d*} \hat{a}'_{\mu} + \text{H.c.} | j_{\text{mp}} \rangle|^2. \quad (108)$$

The aim now is to show these are equivalent. First, we note that in the Coulomb gauge

$$\begin{aligned} \frac{\partial}{\partial t} \hat{\mathcal{A}}_{\perp}(\mathbf{x}) &= \frac{i}{\hbar} [\hat{\mathcal{H}}^C, \hat{\mathcal{A}}_{\perp}(\mathbf{x})] \\ &= -i \sum_{\mu\nu} \frac{\hbar\chi_{\mu\nu}}{\sqrt{2\epsilon_0 \hbar \chi_{\mu\mu}}} \mathbf{f}_{\mu}(\mathbf{x}) \hat{a}_{\nu} + \text{H.c.} \\ &= -\hat{\mathcal{E}}_{\perp}^C(\mathbf{x}), \end{aligned} \quad (109)$$

where  $\hat{\mathcal{E}}_{\perp}^C$  is the correctly truncated transverse electric field operator for the Coulomb gauge defined in Eq. (68). It is generally *not* equal to the transverse electric field operator which one would obtain by simply truncating the expansion in Eq. (99), and this suggests that truncating Eq. (99) directly is the incorrect approach in a dispersive and absorbing medium.

Using Eq. (109), if we take matrix elements with respect to the eigenstates of the Hamiltonian  $\hat{\mathcal{H}}^C$ , we can derive the relationship

$$\langle i_C | \hat{\mathcal{A}}_{\perp} | j_C \rangle = \sum_{\mu\nu} \frac{1}{\sqrt{2\epsilon_0 \hbar \chi_{\mu\mu}}} \frac{\langle i_C | \hbar\chi_{\mu\nu} \mathbf{f}_{\mu} \hat{a}_{\nu} - \text{H.c.} | j_C \rangle}{\omega_j - \omega_i}. \quad (110)$$

Using the fact that  $\omega_d = \omega_j - \omega_i$ , and multiplying by a factor of  $|i\rangle^2 = 1$ , we find

$$R_C = \left| \langle i_C | \left( i \sum_{\mu\nu} \hbar\chi_{\mu\nu}^* \eta_{\nu}^{d*} \hat{a}_{\mu} + \text{H.c.} \right) | j_C \rangle \right|^2. \quad (111)$$

Now, we note that the PZW transformation (excluding the detector subspace), applied to Eq. (111), simply takes  $\hat{a}_{\mu}$  to  $\hat{a}'_{\mu}$ , and the matrix elements to be evaluated in the eigenstates of the multipolar gauge. Thus,  $R_C = R_{\text{mp}}$ , and we can confirm the final result: the operator which should be used to model photodetection at transition frequency  $\omega$  is  $\omega \hat{\mathcal{A}}_{\perp}(\mathbf{x})$ , and this quantity is gauge invariant. Equivalently, the correctly truncated transverse electric field operator  $\hat{\mathcal{E}}_{\perp}^s(\mathbf{x})$  can be used, which is  $\hat{\mathcal{E}}_{\perp}^s = \hat{\mathcal{Y}}_g [\hat{\mathcal{E}}_F]_{\perp} \hat{\mathcal{Y}}_g^{\dagger}$ , and  $[\hat{\mathcal{E}}_F]_{\perp}$  is defined in Eq. (68). This fact can be understood by noting that the matrix element of Eq. (108) is that of the operator  $\mathbf{d}_d \cdot \hat{\mathcal{E}}_{\perp}^{\text{mp}}(\mathbf{r}_d)$ .

We also comment that, as has been pointed out previously, the notion of what the “correct” truncated electric field is can be fundamentally ambiguous, and depends on what is ultimately being measured in experiment [59,62]. In this case, this definition is consistent with photodetection experiments. We also note that this definition of the truncated field violates causality [59]; however, this is perfectly consistent with a truncated mode approximation [78], at least assuming an anharmonic mode spectrum.

## VIII. CONCLUSIONS

In conclusion, we have presented a general theory of gauge-invariant light-matter interactions under material and photonic subspace truncation in an arbitrary medium, focusing on the realistic case of an inhomogeneous dispersive and

absorbing dielectric. Our theory is thus applicable to a wide range of optical and photonic systems.

In Sec. II we gave the fundamental minimal-coupling action which gave Maxwell's equations and the Lorentz force law for a system with medium reservoir oscillator degrees of freedom [ultimately characterized by the dielectric function  $\epsilon(\mathbf{x}, \omega)$ ], as well as free material particles. We then quantized the system using a quantization function approach in Sec. III, where the gauge of the system was left arbitrary, allowing us to show manifestly that the quantum theory was gauge invariant, and easily recover known results for the common Coulomb and multipolar gauges. In Sec. IV we showed how gauge invariance manifests in the quantum theory, and how material and photonic space truncation can be included without sacrificing gauge invariance in the reduced subspace, allowing one to derive models which can be used to model ultrastrong light-matter interactions for common models in quantum optics, including TLSs and single-mode models. We drew contrast with a discrete mode expansion in lossy systems, where the mode expansion Hamiltonian  $\hat{\mathcal{H}}_F$  is nondiagonal, with the more familiar normal mode expansion, which can be implemented in free space, or a lossless and dispersionless dielectric. In particular, we argued that this difference requires rigorous open quantum system models of loss in resonant mode systems (e.g., a cavity-qubit system with photon loss) to be derived in the Coulomb gauge, as this gauge is the unique gauge in which information regarding the part of the “reservoir” electromagnetic field (to be traced out) can be encoded in a purely photonic operator subspace.

Next, in Sec. V we extended the arbitrary-gauge quantization function approach to canonical quantization by allowing for gauge conditions with explicit time dependence, and showed how this could be used to introduce unambiguous phenomenological time-dependent modulation of the transverse light-matter coupling strength. We applied our theory to common models and approximations used in the literature in Sec. VI, including the dipole approximation, the TLS, and how to go beyond the dipole approximation for effective single-particle models. Finally, in Sec. VII we discussed how a particular gauge ambiguity could arise with respect to observables of a mode-truncated field, and how to resolve this ambiguity by means of an explicit “detector particle” model, where gauge invariance is preserved using the techniques of Sec. IV. This allowed us to show that the vector potential is the fundamental field coordinate to be truncated directly, and subsequently, we showed how to obtain the correctly truncated electric field from this result.

Finally, we note that our results can be used as a starting point for more fundamental models of loss and dissipation in quantum optics. For example, the techniques of material and mode truncation can be used to derive rigorous and accurate models of dipole interactions with a few discrete modes, modelled as quantized QNMs (as shown in Appendix C). Separating the entire electromagnetic field into QNM and reservoir components [39,41], and using standard methods of open quantum systems, a master equation could then be derived to govern the lossy dipole-QNM system. This would solve the long-standing problem of how to go beyond a phenomenological formulation of the system-reservoir coupling in quantum optics and input-output theory [52] in the USC

regime for general three-dimensional resonators, which has recently been shown to be essential to predicting emission spectra and other observables from cavity-QED systems in the USC regime [53].

## ACKNOWLEDGEMENTS

We acknowledge funding from Queen's University, the Canadian Foundation for Innovation (CFI), the Natural Sciences and Engineering Research Council of Canada (NSERC), and the Alexander von Humboldt Foundation through a Humboldt Research Award.

## APPENDIX A: ACTION AND EQUATIONS OF MOTION

The total action for the system described in Sec. II is [7,11]

$$\mathcal{S} = \mathcal{S}_{\text{em}}[A_\mu] + \mathcal{S}_{\mathbf{X}}[\mathbf{X}_\omega] + \mathcal{S}_{\mathbf{r}_\alpha}[\mathbf{r}_\alpha] + \mathcal{S}_{\text{int}}[A_\mu; \mathbf{r}_\alpha; \mathbf{X}_\omega], \quad (\text{A1})$$

where

$$\mathcal{S}_{\text{em}}[A_\mu] = \int d^4x \left( \frac{\epsilon_0}{2} \mathbf{E}^2 - \frac{1}{2\mu_0} \mathbf{B}^2 \right), \quad (\text{A2a})$$

$$\mathcal{S}_{\mathbf{r}_\alpha}[\mathbf{r}_\alpha] = \frac{1}{2} \sum_{\alpha} m_{\alpha} \dot{\mathbf{r}}_{\alpha}^2, \quad (\text{A2b})$$

$$\mathcal{S}_{\mathbf{X}}[\mathbf{X}_\omega] = \frac{1}{2} \int d^4x \int_0^{\infty} d\omega (\dot{\mathbf{X}}_{\omega}^2 - \omega^2 \mathbf{X}_{\omega}^2), \quad (\text{A2c})$$

$$\begin{aligned} \mathcal{S}_{\text{int}}[A_\mu; \mathbf{r}_\alpha; \mathbf{X}_\omega] &= \int d^4x \int_0^{\infty} d\omega \alpha(\mathbf{x}, \omega) \mathbf{X}_{\omega} \cdot \mathbf{E} \\ &\quad - \int d^4x J_{\mathbf{A}}^{\mu} A_{\mu}, \end{aligned} \quad (\text{A2d})$$

with

$$\alpha(\mathbf{x}, \omega) = \sqrt{\frac{2\epsilon_0\omega}{\pi}} \epsilon_I(\mathbf{x}, \omega) \quad (\text{A3})$$

and the four-current for the free particles is  $J_{\mathbf{A}}^{\mu}(x) = (c\rho_{\mathbf{A}}(\mathbf{x}, t), \mathbf{J}_{\mathbf{A}}(\mathbf{x}, t))$ . This current satisfies a continuity equation  $\nabla \cdot \mathbf{J}_{\mathbf{A}} + \dot{\rho}_{\mathbf{A}} = 0$ , where

$$\rho_{\mathbf{A}}(\mathbf{x}, t) = \sum_{\alpha} q_{\alpha} \delta(\mathbf{x} - \mathbf{r}_{\alpha}), \quad (\text{A4a})$$

$$\mathbf{J}_{\mathbf{A}}(\mathbf{x}, t) = \sum_{\alpha} q_{\alpha} \dot{\mathbf{r}}_{\alpha} \delta(\mathbf{x} - \mathbf{r}_{\alpha}). \quad (\text{A4b})$$

The electric and magnetic fields are expressed in terms of the potentials as  $\mathbf{E} = -\dot{\mathbf{A}} - \nabla\phi$  and  $\mathbf{B} = \nabla \times \mathbf{A}$ , respectively.

Requiring that the action be stationary, we obtain the following set of equations:

$$\epsilon_0 \nabla \cdot \mathbf{E} + \int_0^{\infty} d\omega \nabla \cdot [\alpha(\mathbf{x}, \omega) \mathbf{X}_{\omega}] = \rho_{\mathbf{A}}, \quad (\text{A5a})$$

$$\frac{1}{\mu_0} \nabla \times \mathbf{B} - \epsilon_0 \dot{\mathbf{E}} - \int_0^{\infty} d\omega \alpha(\mathbf{x}, \omega) \dot{\mathbf{X}}_{\omega} = \mathbf{J}_{\mathbf{A}}, \quad (\text{A5b})$$

$$m_{\alpha} \ddot{\mathbf{r}}_{\alpha} = q_{\alpha} \mathbf{E}(\mathbf{r}_{\alpha}) + q_{\alpha} \dot{\mathbf{r}}_{\alpha} \times \mathbf{B}(\mathbf{r}_{\alpha}), \quad (\text{A5c})$$

$$\ddot{\mathbf{X}}_{\omega} + \omega^2 \mathbf{X}_{\omega} = \alpha(\mathbf{x}, \omega) \mathbf{E}. \quad (\text{A5d})$$

Clearly, Eqs. (A5a) and (A5b) are the Gauss and Ampère-Maxwell laws with active material sources, as well as a

polarization field  $\mathbf{P}_M = \int d\omega \alpha(\mathbf{x}, \omega) \mathbf{X}_\omega$  arising from the passive medium reservoir which also acts as a source—the “passivity” of the medium is demonstrated by the “Fano diagonalization” of Sec. III C, which allows the fields to be expressed in terms of bosonic polariton operators and the photonic Green’s function without explicit reference to the medium interaction. Equation (A5c) is the Lorentz force equation of motion for the active free particles, and Eq. (A5d) is the equation of motion for the passive medium reservoir field.

## APPENDIX B: QUANTIZATION VIA DIRAC’S PRESCRIPTION WITH CONSTRAINTS

In this Appendix, we quantize the system consisting of the electromagnetic fields, passive medium, and active material particles using Dirac’s prescription.

We note that we have three constraints for this system:  $\chi_0 = \chi_1 = \chi_2 = 0$ , as given by Eqs. (3a), (3b), and (9). These constraints are to be understood as applying only on a subset of phase space in which all the constraints are applied—in the parlance of Dirac, these are *weak equalities* [79]. To proceed with the quantization procedure, we must evaluate Poisson brackets between these constraints. In particular, these should be evaluated *before* applying the constraints. It is easy to verify that the Poisson brackets of  $\chi_0$  with the other two constraints vanish. This means  $\chi_0$  is a so-called first-class constraint, and the remaining Poisson brackets (over the “second-class” constraints) thus form an orthogonal matrix  $C_{ij}(\mathbf{x}, \mathbf{x}') = \{\chi_i(\mathbf{x}), \chi_j(\mathbf{x}')\}^1$ , which can be evaluated as

$$\mathbf{C}(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x} - \mathbf{x}') \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (\text{B1})$$

with inverse (in this case)  $\mathbf{C}^{-1} = \mathbf{C}^T$ , which satisfies

$$\sum_k \int d^3x'' C_{ik}(\mathbf{x}, \mathbf{x}'') C_{kj}^{-1}(\mathbf{x}'', \mathbf{x}') = \delta_{ij} \delta(\mathbf{x} - \mathbf{x}'). \quad (\text{B2})$$

From this, one then defines the *Dirac bracket* [12,47,67]:

$$\begin{aligned} \{A, B\}_D &\equiv \{A, B\} \\ &- \sum_{ij} \int d^3x \int d^3x' \{A, \chi_i(\mathbf{x})\} C_{ij}^{-1}(\mathbf{x}, \mathbf{x}') \{\chi_j(\mathbf{x}'), B\}, \end{aligned} \quad (\text{B3})$$

and quantizes by imposing  $[\hat{A}, \hat{B}] = i\hbar\{A, B\}_D$ .

Evaluating in this manner, we find the nonzero commutators to be (in the Schrödinger picture)

$$[\hat{\mathbf{X}}_\omega(\mathbf{x}), \hat{\mathbf{P}}_{\omega'}(\mathbf{x}')] = i\hbar \mathbf{I} \delta(\mathbf{x} - \mathbf{x}') \delta(\omega - \omega'), \quad (\text{B4})$$

$$[\hat{\mathbf{r}}_\alpha, \hat{\mathbf{p}}_{\alpha'}] = i\hbar \mathbf{I} \delta_{\alpha\alpha'}, \quad (\text{B5})$$

$$\begin{aligned} [\hat{\mathbf{A}}(\mathbf{x}), \hat{\mathbf{P}}_A(\mathbf{x}')] &= i\hbar [\mathbf{I} \delta(\mathbf{x} - \mathbf{x}') + \nabla^x \mathbf{K}(\mathbf{x}', \mathbf{x})] \\ &= i\hbar [\delta^\perp(\mathbf{x} - \mathbf{x}') + \nabla^x \mathbf{K}_\perp(\mathbf{x}', \mathbf{x})], \end{aligned} \quad (\text{B6})$$

<sup>1</sup>Here, for example, the Poisson bracket for fields  $f(\mathbf{x})$  and  $g(\mathbf{x}')$  which depend on fields  $\theta_i$  and their conjugates  $\Pi_{\theta_i}$  takes the form  $\{f(\mathbf{x}), g(\mathbf{x}')\} = \sum_i \int d^3y [\frac{\delta f(\mathbf{x})}{\delta \theta_i(\mathbf{y})} \frac{\delta g(\mathbf{x}')}{\delta \Pi_{\theta_i}(\mathbf{y})} - \frac{\delta g(\mathbf{x}')}{\delta \theta_i(\mathbf{y})} \frac{\delta f(\mathbf{x})}{\delta \Pi_{\theta_i}(\mathbf{y})}]$ , with all fields evaluated at equal times.

and

$$[\hat{\mathbf{p}}_\alpha, \hat{\mathbf{P}}_A(\mathbf{x})] = i\hbar q_\alpha \nabla^{\hat{\mathbf{r}}_\alpha} \mathbf{K}(\mathbf{x}, \hat{\mathbf{r}}_\alpha). \quad (\text{B7})$$

We can also construct the quantum Hamiltonian:

$$\begin{aligned} \hat{H} &= \sum_\alpha \hat{\mathbf{r}}_\alpha \cdot \hat{\mathbf{p}}_\alpha + \int d^3x \hat{\mathbf{A}} \cdot \hat{\mathbf{P}}_A \\ &+ \int d^3x \int_0^\infty d\omega \hat{\mathbf{X}}_\omega \cdot \hat{\mathbf{P}}_{\mathbf{X}_\omega} - \hat{\mathcal{L}}. \end{aligned} \quad (\text{B8})$$

While these commutators define, in principle, a working quantum theory of fields, they do not separate photonic and material degrees of freedom in the form of canonical commutation relations; that is, the operators  $\hat{\mathbf{A}}$  and  $\hat{\mathbf{P}}_A$  cannot be expanded as a superposition of bosonic creation and annihilation operators. Nonetheless, a canonical transformation can be easily found to describe the system in terms of new coordinates which do satisfy the usual canonical commutation relations.

After quantization, the “weak equalities” of the constraints become strong equalities, and can be applied to the quantum operators. The constraint  $\chi_1(\mathbf{x}) = 0$  means that the longitudinal part of  $\hat{\mathbf{P}}_A(\mathbf{x})$  can be expressed analytically in terms of  $\hat{\rho}_A(\mathbf{x})$ :

$$\begin{aligned} \hat{\mathbf{P}}_{A,\parallel}(\mathbf{x}) &= \hat{\mathbf{P}}_\parallel(\mathbf{x}) \\ &= \sum_\alpha q_\alpha (\hat{\mathbf{r}}_\alpha - \mathbf{r}_A) \cdot \int_0^1 ds \delta^\parallel[\mathbf{x} - \mathbf{r}_A - s(\hat{\mathbf{r}}_\alpha - \mathbf{r}_A)], \end{aligned} \quad (\text{B9})$$

where the second equality can be shown to be consistent with Eq. (12). For this work, we take  $\mathbf{r}_A = \mathbf{0}$  hereafter without loss of generality. Note that, more generally, one can also consider situations where  $\mathbf{r}_A \rightarrow \hat{\mathbf{r}}_A$  corresponds to a dynamical degree of freedom (e.g., a center-of-mass coordinate) [18,80].

By means of a canonical transformation, we can define new *unconstrained* canonical coordinates  $\hat{\mathbf{A}}_\perp$  and  $\hat{\mathbf{P}} = \hat{\mathbf{P}}_A - \hat{\mathbf{P}}$  [12]. In this new coordinate system, the field variables are manifestly transverse, and the Gauss’s law constraint simply becomes  $\nabla \cdot \hat{\mathbf{P}} = 0$ . One can then verify that we have the modified commutators from Sec. III B, which now take their usual canonical form.

## APPENDIX C: CONSTRUCTION OF DISCRETE MODES FROM THE BOSONIC CONTINUUM AND RELATIONSHIP TO QUANTIZED QUASINORMAL MODE THEORY

In this Appendix, we first detail in Sec. C 1 how discrete modes can be constructed out of the continuous mode expansion expressed in terms of the macroscopic QED operators  $\hat{\mathbf{b}}(\mathbf{r}, \omega)$ ,  $\hat{\mathbf{b}}^\dagger(\mathbf{r}, \omega)$ , as was used to discuss mode truncation in Sec. IV C, and then in Sec. C 2 connect to the important example of quantized QNMs.

### 1. Construction of discrete modes in system with dispersion and absorption

In general, we can construct a discrete “modal” operator as  $\hat{a}_\mu = \int d^3x \int_0^\infty \mathbf{L}_\mu(\mathbf{x}, \omega) \cdot \hat{\mathbf{b}}(\mathbf{x}, \omega)$ , where  $\mathbf{L}_\mu(\mathbf{x}, \omega)$  is a function which projects the full bosonic subspace of the Fano

diagonalization onto a discrete Fock subspace. For this to be true, we must require  $[\hat{a}_\mu, \hat{a}_\eta^\dagger] = \delta_{\mu\eta}$ , and so

$$\int d^3x \int_0^\infty d\omega \mathbf{L}_\mu(\mathbf{x}, \omega) \cdot \mathbf{L}_\eta^*(\mathbf{x}, \omega) = \delta_{\mu\eta}. \quad (\text{C1})$$

Under this definition, we have  $\hat{\mathcal{H}}_F = \sum_{\mu\eta} \hbar \chi_{\mu\eta} \hat{a}_\mu^\dagger \hat{a}_\eta$ , where

$$\chi_{\mu\eta} = \int d^3x \int_0^\infty d\omega \omega \mathbf{L}_\mu(\mathbf{x}, \omega) \cdot \mathbf{L}_\eta^*(\mathbf{x}, \omega), \quad (\text{C2})$$

$$\hat{P}_M \hat{\mathbf{b}}(\mathbf{x}, \omega) \hat{P}_M = \sum_\mu \mathbf{L}_\mu^*(\mathbf{x}, \omega) \hat{a}_\mu, \quad (\text{C3})$$

and

$$\begin{aligned} \mathbf{f}_\mu(\mathbf{x}) &= \sqrt{\frac{2\chi_{\mu\mu}}{\pi}} \int d^3x' \int_0^\infty \frac{d\omega}{\omega} \sqrt{\epsilon_I(\mathbf{x}', \omega)} \mathbf{G}^\perp(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{L}_\mu^*(\mathbf{x}', \omega) \\ & \quad (\text{C4}) \end{aligned}$$

are the mode profiles for the expansion of the vector potential in Eq. (61). Alternatively, one could define mode profiles for an expansion of the transverse electric field, as in Eq. (99):

$$\begin{aligned} \mathbf{f}_{\mu^\perp}^{\mathbf{E}}(\mathbf{x}) &= \sqrt{\frac{2}{\chi_{\mu\mu}\pi}} \int d^3x' \int_0^\infty d\omega \sqrt{\epsilon_I(\mathbf{x}', \omega)} \mathbf{G}^\perp(\mathbf{x}, \mathbf{x}', \omega) \cdot \mathbf{L}_\mu^*(\mathbf{x}', \omega). \\ & \quad (\text{C5}) \end{aligned}$$

However, to preserve gauge invariance under truncation, the properly truncated electric field operator should instead be expanded in terms of the functions  $\mathbf{f}'_\mu = \sum_\eta \frac{\chi_{\mu\eta}^*}{\sqrt{\chi_{\mu\mu}\chi_{\eta\eta}}} \mathbf{f}_\eta$ .

This construction is general, but can be used to describe, e.g., quantized QNMs [39]. Additionally, these “modes” need not be true modes in the sense that they satisfy Maxwell’s equations with the appropriate boundary conditions—rather, they constitute, in the most general case, a truncation of the spatial and frequency-dependent degrees of freedom of the “Fano-diagonalized” polariton operators in which the electromagnetic and passive medium fields are expressed as a linear functional thereof.

By appropriate choice of projection functions, the untruncated results can be recovered by imposing an appropriate completeness relation. Specifically, if the projection functions satisfy

$$\sum_\mu \mathbf{L}_\mu(\mathbf{x}, \omega) \mathbf{L}_\mu^*(\mathbf{x}', \omega') = \mathbf{I} \delta(\mathbf{x} - \mathbf{x}') \delta(\omega - \omega'), \quad (\text{C6})$$

then

$$\begin{aligned} & \sum_{\mu\eta} \frac{\chi_{\mu\eta}}{\sqrt{\chi_{\mu\mu}\chi_{\eta\eta}}} \mathbf{f}_\mu(\mathbf{x}) \mathbf{f}_\eta^*(\mathbf{x}') \\ &= \frac{2}{\pi} \int d^3y \int_0^\infty \frac{d\omega}{\omega} \epsilon_I(\mathbf{y}, \omega) \mathbf{G}^\perp(\mathbf{x}, \mathbf{y}, \omega) \cdot \mathbf{G}^{\perp,*}(\mathbf{y}, \mathbf{x}', \omega) \\ &= \frac{2}{\pi} \int_0^\infty \frac{d\omega}{\omega} \text{Im}[\mathbf{G}^\perp(\mathbf{x}, \mathbf{x}', \omega)] \\ &= \delta^\perp(\mathbf{x} - \mathbf{x}'), \end{aligned} \quad (\text{C7})$$

where in the second line we have used Eqs. (C2) and (C6), in the third the relations (32), (33), and (C2), and in the fourth an identity related to causality proven in Ref. [5].

Similarly one can show  $\sum_\mu \mathbf{f}_\mu^{\mathbf{E}\perp}(\mathbf{x}) \mathbf{f}_\mu(\mathbf{x}') = \delta^\perp(\mathbf{x} - \mathbf{x}')$ . Note that the completeness relation can be put in the more usual form by introducing symmetrized mode functions  $\mathbf{f}_\mu^s(\mathbf{x}) = \sum_\eta [\chi^{\frac{1}{2}}]_{\eta\mu} \mathbf{f}_\eta(\mathbf{x}) / \sqrt{\chi_{\eta\eta}}$ . That the matrix  $\chi^{\frac{1}{2}}$  necessarily exists relies on  $\chi$  being invertible. From the form of Eq. (C2), and the orthogonality and completeness relations imposed on the projection functions  $\mathbf{L}_\mu(\mathbf{x}, \omega)$ ,  $\chi$  takes the form of an inner product with positive weight function  $\omega$ , and thus is invertible. The symmetrized mode functions then satisfy  $\sum_\mu \mathbf{f}_\mu^s(\mathbf{x}) \mathbf{f}_\mu^{s,*}(\mathbf{x}') = \delta^\perp(\mathbf{x} - \mathbf{x}')$ .

For an example, a mode expansion that is complete in the spatial part can be found by using plane waves as the basis set for  $\mathbf{L}$ . In this case the spatial part could take the form  $(\frac{k_0}{2\pi})^{\frac{3}{2}} e^{i\mathbf{k}_n \cdot \mathbf{x}}$  where  $\mathbf{k}_n = (n_x, n_y, n_z)k_0$ , with  $n_x, n_y, n_z$  integers, and  $k_0$  being a momentum cutoff which would go to zero as the quantization volume  $\rightarrow \infty$ .

It should be noted that this expansion is not the same as the common *normal mode* approach often taken in dielectric media without dispersion or absorption, where the mode functions are solutions to the Helmholtz equation. In the absence of loss or dispersion, and with closed or periodic boundary conditions, these are the generalized eigenfunctions which satisfy  $\int d^3x \epsilon(\mathbf{x}) \mathbf{h}_i(\mathbf{x}) \cdot \mathbf{h}_j(\mathbf{x}) = \delta_{ij}$ . These functions are *generalized transverse*, in that they satisfy  $\nabla \cdot [\epsilon(\mathbf{x}) \mathbf{h}_i(\mathbf{x})] = 0$ . Only in the limit  $\epsilon_R(\mathbf{x}, \omega) \rightarrow \epsilon(\mathbf{x})$  and  $\epsilon_I(\mathbf{x}, \omega) \rightarrow 0$ , and for closed or periodic boundary conditions, can one expand the fields as linear combinations of these generalized transverse eigenfunctions. In fact, our approach does not even require the modal functions  $\mathbf{f}_\mu(\mathbf{r})$  to satisfy the Helmholtz equation.

In the case that the projection functions  $\mathbf{L}_\mu(\mathbf{x}, \omega)$  are well isolated in frequency, one can approximate the matrix  $\chi_{\mu\eta}$  in Eq. (C2) as diagonal, with  $\omega_\mu \equiv \chi_{\mu\mu}$ , and in this case the generalized completeness relation, as well as the mode expansions take their more usual forms  $\hat{\mathcal{A}}_\perp(\mathbf{x}) \approx \sum_\mu \sqrt{\frac{\hbar}{2\epsilon_0\omega_\mu}} \mathbf{f}(\mathbf{x}) \hat{a}_\mu + \text{H.c.}$ , and  $[\hat{\mathcal{E}}_F(\mathbf{x})]_\perp \approx i \sum_\mu \sqrt{\frac{\hbar\omega_\mu}{2\epsilon_0}} \mathbf{f}(\mathbf{x}) \hat{a}_\mu + \text{H.c.}$

As shown in Appendix D, the development of the arbitrary-gauge quantization of the electromagnetic field is implemented more naturally for lossless and dispersionless systems with closed or periodic boundary conditions by utilizing a constraint equation [analogous to  $\chi_2$ —Eq. (9)] more suitable for the implementation of the generalized Coulomb

and multipolar gauges, as well as an appropriate modified Lagrangian. Such an approach yields mode expansions directly in terms of the generalized eigenfunctions to the Helmholtz equations.

## 2. Relationship to quantized quasinormal modes

A particularly noteworthy example of the construction of discrete modes out of the bosonic continuum is in the quantized quasinormal mode (QNM) theory [39,42,81]; QNMs, with mode profiles  $\tilde{\mathbf{f}}_\mu(\mathbf{x})$ , are eigenfunctions of the Helmholtz equation, which for open boundary conditions is a non-Hermitian eigenproblem:

$$\nabla \times \nabla \times \tilde{\mathbf{f}}_\mu(\mathbf{x}) - \frac{\tilde{\omega}_\mu^2}{c^2} \epsilon(\mathbf{x}, \tilde{\omega}_\mu) \tilde{\mathbf{f}}_\mu(\mathbf{x}) = 0, \quad (\text{C8})$$

along with suitable boundary conditions [i.e., the Silver-Müller conditions of Eq. (29)]. The eigenfrequency  $\tilde{\omega}_\mu = \omega_\mu - i\gamma_\mu$  is the complex QNM frequency, with real part  $\omega_\mu$  and imaginary (dissipative) part  $\gamma_\mu$ ; QNMs allow one to calculate relevant quantum optics quantities such as cavity  $Q$  factors, complex effective mode volumes, Purcell factors, and radiative  $\beta$  factors [10,17,37,38]. Notably, QNMs are the open-cavity modes for both dielectric resonators and systems with material loss (e.g., also describing localized plasmon modes of metallic resonators, which are also QNMs, with radiative and nonradiative loss). In the case of lossless dielectrics (i.e., with no material loss), the dissipation captured in  $\gamma_\mu$  is entirely through radiative loss.

For quantized QNMs, the mode projection functions  $\mathbf{L}_\mu(\mathbf{x}, \omega)$  take the form

$$\mathbf{L}_\mu(\mathbf{x}, \omega) = \sum_\eta [S^{-\frac{1}{2}}]_{\mu\eta} \sqrt{\frac{\omega_\eta}{2\pi}} \frac{\sqrt{\epsilon_I(\mathbf{x}, \omega)}}{\tilde{\omega}_\eta - \omega} \tilde{\mathbf{F}}_\eta^I(\mathbf{x}, \omega). \quad (\text{C9})$$

While these are orthonormal, they are not complete, and so approximating the field expansion in terms of purely QNM operators is necessarily a form of mode truncation, which requires the procedures outlined in Sec. IV C to restore gauge invariance. The symmetrizing matrix  $S_{\mu\eta}$  is given by the frequency integral

$$S_{\mu\eta} = \frac{\sqrt{\omega_\mu \omega_\eta}}{2\pi} \int_0^\infty d\omega \frac{S_{\mu\eta}^{\text{rad}}(\omega) + S_{\mu\eta}^{\text{grad}}(\omega)}{(\omega - \tilde{\omega}_\mu)(\omega - \tilde{\omega}_\eta^*)}, \quad (\text{C10})$$

where

$$S_{\mu\eta}^{\text{rad}}(\omega) = \int_V d^3x \epsilon_I(\mathbf{x}, \omega) \tilde{\mathbf{f}}_\mu(\mathbf{x}) \cdot \tilde{\mathbf{f}}_\eta^*(\mathbf{x}) \quad (\text{C11})$$

gives the nonradiative loss contribution through a QNM overlap integral over a region  $V$  containing the absorptive material, which is assumed to be separated from the rest of space by a discontinuity in the dielectric function, and

$$\begin{aligned} S_{\mu\eta}^{\text{grad}}(\omega) &= \int_{\mathbb{R}^3 - V} d^3x \epsilon_I(\mathbf{x}, \omega) \tilde{\mathbf{F}}_\mu(\mathbf{x}, \omega) \cdot \tilde{\mathbf{F}}_\eta^*(\mathbf{x}, \omega) \\ &= \frac{n_{\text{BC}}}{\omega} \oint_{S_\infty} dA_s \tilde{\mathbf{F}}_\mu(\mathbf{s}, \omega) \cdot \tilde{\mathbf{F}}_\eta^*(\mathbf{s}, \omega) \end{aligned} \quad (\text{C12})$$

gives the radiative loss through a far-field surface  $S_\infty$ .

The function  $\tilde{\mathbf{F}}'_\mu(\mathbf{x}, \omega)$ , in Eq. (C9), is defined piecewise, where within the absorptive region  $V$  it is equal to  $\tilde{\mathbf{f}}_\mu(\mathbf{x})$ , and

outside of this region it is defined as  $\tilde{\mathbf{F}}_\mu(\mathbf{x}, \omega)$ , which is a regularized form to avoid the divergence of  $\tilde{\mathbf{f}}_\mu(\mathbf{x})$  in the far field and can be obtained from a Dyson equation approach [41], or a near-field to far-field transformation [82].

The transverse Green's function for locations  $\mathbf{x}$ , within the volume  $V$ , can be expanded as [83]

$$\mathbf{G}^\perp(\mathbf{x}, \mathbf{x}', \omega) = \sum_\mu \frac{\omega}{2(\tilde{\omega}_\mu - \omega)} \tilde{\mathbf{f}}(\mathbf{x}) \tilde{\mathbf{F}}'_\mu(\mathbf{x}', \omega). \quad (\text{C13})$$

Substituting these results into Eq. (C4), we find, for  $\mathbf{x}$  within the volume  $V$ ,

$$\mathbf{f}_\mu(\mathbf{x}) = \sum_\eta \tilde{\mathbf{f}}_\eta(\mathbf{x}) [S^{\frac{1}{2}}]_{\eta\mu} \sqrt{\frac{\chi_{\mu\mu}}{\omega_\eta}}. \quad (\text{C14})$$

This expression also holds approximately for locations near the volume  $V$ , e.g., in plasmonic systems where the field is dominated by the QNMs in the near vicinity of the scattering structure [83,84].

The functions  $\mathbf{f}_\mu(\mathbf{x})$  are similar to symmetrized QNM mode functions which have appeared in previous works. However, previous works have performed the mode expansion in terms of the electric field, as opposed to the vector potential (which plays a more fundamental role in ultrastrong light-matter interactions), so the expressions we show here are slightly different.

As a consistency check, note that due to the pole in Eq. (C9), for  $Q_\mu \equiv \omega_\mu / (2\gamma_\mu) \gg 1$ , the frequency integral in the definition of  $\chi_{\mu\mu}$  is sharply peaked about  $\omega = \omega_\mu$  and thus we can approximate  $\chi_{\mu\mu} \approx \omega_\mu$ . In this common case, we can also expand  $[\hat{\mathcal{E}}_F(\mathbf{x})]_\perp \approx i \sum_\mu \sqrt{\frac{\hbar\omega_\mu}{2\epsilon_0}} \mathbf{f}_\mu(\mathbf{x}) \hat{a}_\mu + \text{H.c.}$ , which is the same result reached in previous works on quantized QNMs [39].

## APPENDIX D: QUANTIZATION AND TRUNCATION IN A DIELECTRIC MEDIUM WITHOUT DISPERSION OR DISSIPATION

For a medium with no dispersion or absorption, the dielectric function is real and independent of frequency such that  $\epsilon = \epsilon(\mathbf{x})$ . In this case Maxwell's equations yield eigenmodes of the medium,  $\mathbf{h}_\mu(\mathbf{x})$ , which satisfy the Helmholtz equation,

$$\nabla \times \nabla \times \mathbf{h}_\mu(\mathbf{x}) - \frac{\omega_\mu^2}{c^2} \epsilon(\mathbf{x}) \mathbf{h}_\mu(\mathbf{x}) = 0, \quad (\text{D1})$$

and when solved with closed or periodic boundary conditions, yields corresponding (real) eigenvalues  $\omega_\mu$ . The eigenfunctions are generalized transverse, in that they satisfy  $\nabla \cdot [\epsilon(\mathbf{x}) \mathbf{h}_\mu(\mathbf{x})] = 0$  and can be chosen to be real [65]. These are *normal modes*, which are orthogonal in the sense that  $\int d^3x \epsilon(\mathbf{x}) \mathbf{h}_\mu(\mathbf{x}) \cdot \mathbf{h}_\nu(\mathbf{x}) = \delta_{\mu\nu}$ , and complete in the sense that any generalized transverse function can be expanded as a linear combination of them [63,65,85]. In Sec. D 1 we perform canonical quantization using Dirac's constrained quantization in the generalized Coulomb gauge, and show how the fields can be expanded in terms of the generalized transverse normal mode expansion; then in Sec. D 2 we perform a PZW transformation to obtain the result for the generalized multipolar gauge, recovering previously known results [63–65,85].

In Sec. D3 we truncate the material and mode degrees of freedom for this system in a manner consistent with gauge invariance in the truncated space.

Note for open resonators and cavities, even with no material loss (i.e., a real dielectric), one should formally use QNMs if performing few mode quantization (as desired in cavity-QED), though mode quantization using normal modes with heuristic dissipation is often a very good approximation for single-mode high- $Q$  resonators outside of the USC regime. However, with several cavity modes, it can be essential to capture the effects of the QNM phase [39,86,87], even when not considering USC. Performing quantization with material loss and carefully taking the limit of a lossless dielectric naturally recovers a description that is valid for the QNMs of dielectric cavity systems, which is also consistent with the fluctuation-dissipation theorem [81]. As noted in Appendix C, in that case, the dissipation is fully contained through the *radiative* loss, which is an inherent part of the QNM.

### 1. Canonical quantization in the generalized Coulomb gauge

While results for a nonabsorbing and nondispersive dielectric can be obtained as a limiting case of the theory presented in the main text [81], it is more convenient to start with a different effective Lagrangian which does not explicitly contain the medium reservoir degrees of freedom, and gives the macroscopic Maxwell's equations upon extremization of the action:

$$\begin{aligned} \mathcal{L}_r = & \frac{1}{2} \sum_{\alpha} m_{\alpha} \dot{\mathbf{r}}_{\alpha}^2 + \frac{1}{2} \int d^3x \left[ \epsilon_0 \epsilon(\mathbf{x}) \mathbf{E}^2 - \frac{1}{\mu_0} \mathbf{B}^2 \right] \\ & - \int d^3x J_A^{\mu} A_{\mu}. \end{aligned} \quad (\text{D2})$$

This Lagrangian gives the same canonical momenta as the one in the main text for the scalar potential and particle coordinates:  $\Pi_{\phi} = 0$ ,  $\mathbf{p}_{\alpha} = m_{\alpha} \dot{\mathbf{r}}_{\alpha} + q_{\alpha} \mathbf{A}(\mathbf{r}_{\alpha})$ , and gives for the vector potential the canonical momentum  $\Pi_{\mathbf{A}} = -\epsilon_0 \epsilon(\mathbf{x}) \mathbf{E}$ . The constraint  $\chi_0 = \Pi_{\phi} = 0$  is thus unchanged, as well as Gauss's law  $\chi_1 = \nabla \cdot \Pi_{\mathbf{A}} - \rho_A = 0$ , provided it is expressed in terms of the new canonical momentum.

To quantize this system in the generalized Coulomb gauge, one approach is to use the quantization function  $\mathbf{K}(\mathbf{x}, \mathbf{x}')$  as in Sec. III B, where the transverse part of this function determines the gauge. We briefly mention how this can be done later, with appropriate constraints on the form of  $\mathbf{K}$ . Alternatively, we can take a more direct approach, explicitly let the gauge constraint be  $\chi_2 = \nabla \cdot (\epsilon \mathbf{A}) = 0$ , and calculate the Dirac brackets from this. Choosing the latter approach, the nonzero elements of the constraint matrix are  $C_{12}(\mathbf{x}, \mathbf{x}')$  and  $C_{21}(\mathbf{x}, \mathbf{x}') = -C_{12}(\mathbf{x}', \mathbf{x})$ , where

$$C_{12}(\mathbf{x}, \mathbf{x}') = \nabla^{\mathbf{x}} \cdot [\epsilon(\mathbf{x}) \nabla^{\mathbf{x}} \delta(\mathbf{x} - \mathbf{x}')]. \quad (\text{D3})$$

This implies

$$-\nabla^{\mathbf{x}} \cdot [\epsilon(\mathbf{x}) \nabla^{\mathbf{x}} C_{12}^{-1}(\mathbf{x}, \mathbf{x}')] = \delta(\mathbf{x} - \mathbf{x}'). \quad (\text{D4})$$

While we do not solve for an explicit form for the inverse constraint matrix, that  $C_{12}^{-1}$  satisfies Eq. (D4)

is in fact the only information needed to quantize the theory.

Constructing the Dirac bracket between the field coordinate and momentum, we obtain

$$\begin{aligned} Q_{ij}(\mathbf{x}, \mathbf{x}') & \equiv \{A_i(\mathbf{x}), \Pi_{A_j}(\mathbf{x}')\}_{\text{D}} \\ & = \delta_{ij} \delta(\mathbf{x} - \mathbf{x}') - \epsilon(\mathbf{x}') \nabla_j^{\mathbf{x}'} \nabla_i^{\mathbf{x}} C_{12}^{-1}(\mathbf{x}, \mathbf{x}'). \end{aligned} \quad (\text{D5})$$

We can show that this function is generalized transverse in its first argument, in that it satisfies

$$\nabla_i^{\mathbf{x}} [\epsilon(\mathbf{x}) Q_{ij}(\mathbf{x}, \mathbf{x}')] = 0. \quad (\text{D6})$$

To prove this, we use Eq. (D4), obtaining

$$\begin{aligned} \nabla_i^{\mathbf{x}} [\epsilon(\mathbf{x}) Q_{ij}(\mathbf{x}, \mathbf{x}')] \\ = \nabla_j^{\mathbf{x}} [\epsilon(\mathbf{x}) \delta(\mathbf{x} - \mathbf{x}')] + \epsilon(\mathbf{x}') \nabla_j^{\mathbf{x}'} \delta(\mathbf{x} - \mathbf{x}'), \end{aligned} \quad (\text{D7})$$

and to show this is zero, we can integrate against a test function:

$$\int d^3x' \nabla_i^{\mathbf{x}} [\epsilon(\mathbf{x}) Q_{ij}(\mathbf{x}, \mathbf{x}')] f_j(\mathbf{x}') = 0. \quad (\text{D8})$$

Similarly, one can show that  $\nabla_j^{\mathbf{x}'} Q_{ij}(\mathbf{x}, \mathbf{x}') = 0$ , i.e., it is transverse with respect to its second argument. Since  $Q_{ij}(\mathbf{x}, \mathbf{x}')$  is generalized transverse with respect to  $\mathbf{x}$  and  $i$ , it must have an expansion within the generalized eigenfunctions of the Helmholtz equation:  $Q_{ij}(\mathbf{x}, \mathbf{x}') = \sum_{\mu} Q_{\mu,j}(\mathbf{x}') h_{\mu,i}(\mathbf{x})$ . The expansion coefficients  $Q_{\mu,j}(\mathbf{x}')$  can be found as  $\int d^3x \mathbf{h}_{\mu}(\mathbf{x}) \cdot \mathbf{Q}(\mathbf{x}, \mathbf{x}')$ , and so

$$\begin{aligned} Q_{\mu,j}(\mathbf{x}') & = \epsilon(\mathbf{x}') h_{\mu,j}(\mathbf{x}') \\ & - \epsilon(\mathbf{x}') \nabla_j^{\mathbf{x}'} \int d^3x \epsilon(\mathbf{x}) h_{\mu,i}(\mathbf{x}) \nabla_i^{\mathbf{x}} C_{12}^{-1}(\mathbf{x}, \mathbf{x}'). \end{aligned} \quad (\text{D9})$$

Since the integral in the second line of Eq. (D9) consists of a transverse function  $\epsilon \mathbf{h}_{\mu}$  integrated against a longitudinal function  $\nabla C_{12}^{-1}$ , it vanishes. As such, we can write

$$Q_{ij}(\mathbf{x}, \mathbf{x}') = \sum_{\mu} \epsilon(\mathbf{x}') h_{\mu,i}(\mathbf{x}) h_{\mu,j}(\mathbf{x}'). \quad (\text{D10})$$

In this form, it can be seen that  $\mathbf{Q}$  is the generalized transverse delta function, introduced in other works [64,65]. For any vector function  $\mathbf{f}(\mathbf{x})$ , one can define

$$\begin{aligned} [\hat{\mathbf{f}}(\mathbf{x})]_{\perp}^{(\epsilon)} & = \int d^3x' \mathbf{Q}(\mathbf{x}, \mathbf{x}') \cdot \mathbf{f}(\mathbf{x}') \\ & = \sum_{\mu} \mathbf{h}_{\mu}(\mathbf{x}) \int d^3x' \epsilon(\mathbf{x}') \mathbf{h}_{\mu}(\mathbf{x}') \cdot \mathbf{f}(\mathbf{x}'), \end{aligned} \quad (\text{D11})$$

where  $[\mathbf{f}(\mathbf{x})]_{\perp}^{(\epsilon)} = [\mathbf{f}_{\perp}(\mathbf{x})]_{\perp}^{(\epsilon)}$  denotes the generalized transverse component of  $\mathbf{f}$ , which satisfies

$$\nabla \cdot (\epsilon(\mathbf{x}) [\hat{\mathbf{f}}(\mathbf{x})]_{\perp}^{(\epsilon)}) = 0. \quad (\text{D12})$$

Now, the only other nonzero Dirac brackets are  $\{r_{\alpha,i}, p_{\alpha',j}\}_{\text{D}} = \delta_{\alpha\alpha'} \delta_{ij}$ , and

$$\{p_{\alpha,i}, \Pi_{A_j}(\mathbf{x})\}_{\text{D}} = -q_{\alpha} \epsilon(\mathbf{x}) \nabla_i^{\mathbf{r}_{\alpha}} \nabla_j^{\mathbf{x}} C_{12}^{-1}(\mathbf{r}_{\alpha}, \mathbf{x}). \quad (\text{D13})$$

Thus, we can construct separable photonic and matter Hilbert spaces by, similarly as in Sec. III B, defining a new canonical

variable:  $\mathbf{\Pi} = \mathbf{\Pi}_A - \mathbf{P}^{\text{gC}}$ , where

$$\mathbf{P}^{\text{gC}}(\mathbf{x}) = - \int d^3x' \mathbf{F}(\mathbf{x}, \mathbf{x}') \rho_A(\mathbf{x}), \quad (\text{D14})$$

and  $\mathbf{F}(\mathbf{x}, \mathbf{x}') = -\epsilon(\mathbf{x}) \nabla^x C_{12}^{-1}(\mathbf{x}', \mathbf{x})$ . This function is defined by two important properties:  $\mathbf{F}(\mathbf{x}, \mathbf{x}')/\epsilon(\mathbf{x})$  is longitudinal, and its longitudinal part is the Green's function for the divergence operator; i.e.,  $\mathbf{F}_{\parallel}(\mathbf{x}, \mathbf{x}') = \mathbf{K}_{\parallel}(\mathbf{x}, \mathbf{x}')$ . This latter property ensures that  $\mathbf{\Pi}$  is transverse, as can be seen from the constraint  $\chi_1$  (because  $\nabla \cdot \mathbf{P}^{\text{gC}} = -\rho_A$ ). Note that one could alternatively arrive at the quantized theory in the generalized Coulomb gauge by using a quantization function  $\mathbf{K}$  which satisfies these same properties as  $\mathbf{F}$  as the quantization function in the approach from Sec. III B.

We can now promote Dirac brackets (multiplied by  $i\hbar$ ) to commutators in the usual fashion, and promote canonical variables to operators. We thus have

$$[\hat{\mathbf{A}}(\mathbf{x}), \hat{\mathbf{\Pi}}(\mathbf{x}')] = i\hbar \mathbf{Q}(\mathbf{x}, \mathbf{x}'), \quad (\text{D15})$$

which implies that the fields have modal expansions [63]

$$\hat{\mathbf{A}}(\mathbf{x}) = \sum_{\mu} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mu}}} \mathbf{h}_{\mu}(\mathbf{x}) \hat{a}_{\mu} + \text{H.c.}, \quad (\text{D16a})$$

$$\hat{\mathbf{\Pi}}(\mathbf{x}) = -i\epsilon_0\epsilon(\mathbf{x}) \sum_{\mu} \sqrt{\frac{\hbar\omega_{\mu}}{2\epsilon_0}} \mathbf{h}_{\mu}(\mathbf{x}) \hat{a}_{\mu} + \text{H.c.} \quad (\text{D16b})$$

We then construct the Hamiltonian in the usual manner, finding

$$\begin{aligned} \hat{H} &= \sum_{\alpha} \frac{[\hat{\mathbf{p}}_{\alpha} - q_{\alpha} \hat{\mathbf{A}}(\hat{\mathbf{r}}_{\alpha})]^2}{2m_{\alpha}} \\ &+ \frac{1}{2} \left[ \int d^3x \frac{(\hat{\mathbf{\Pi}} + \hat{\mathbf{P}}^{\text{gC}})^2}{\epsilon_0\epsilon(\mathbf{x})} + \frac{\hat{\mathbf{B}}^2}{\mu_0} \right] - \int d^3x \hat{\mathbf{P}}^{\text{gC}} \cdot \nabla \hat{\phi}. \end{aligned} \quad (\text{D17})$$

We can now make several simplifications. First, we note that the term  $\int \hat{\mathbf{\Pi}} \cdot \hat{\mathbf{P}}^{\text{gC}}/\epsilon(\mathbf{x})$  vanishes, because  $\hat{\mathbf{\Pi}}$  is transverse, and  $\hat{\mathbf{P}}^{\text{gC}}/\epsilon(\mathbf{x})$  is longitudinal. Second, the last term can be written as

$$\int d^3x \hat{\mathbf{P}}^{\text{gC}} \cdot \nabla \hat{\phi} = \int d^3x \frac{\hat{\mathbf{P}}^{\text{gC}}}{\epsilon(\mathbf{x})} [\epsilon(\mathbf{x}) \nabla \hat{\phi}]_{\parallel}, \quad (\text{D18})$$

but the longitudinal part of  $(\epsilon(\mathbf{x}) \nabla \hat{\phi})$  can be found from the constraint  $\chi_1$ , by noting that since  $\epsilon(\mathbf{x}) \hat{\mathbf{A}}$  is transverse; then the constraint simply becomes

$$\nabla \cdot [\epsilon_0\epsilon(\mathbf{x}) \nabla \hat{\phi}] = -\hat{\rho}_A = \nabla \cdot \hat{\mathbf{P}}^{\text{gC}}. \quad (\text{D19})$$

Thus, the total Hamiltonian in the generalized Coulomb gauge  $\hat{H}^{\text{gC}}$  can be written

$$\hat{H}^{\text{gC}} = \sum_{\alpha} \frac{[\hat{\mathbf{p}}_{\alpha} - q_{\alpha} \hat{\mathbf{A}}(\hat{\mathbf{r}}_{\alpha})]^2}{2m_{\alpha}} + \hat{H}_{\text{F}} + \hat{V}_{\text{Coul}}, \quad (\text{D20})$$

with

$$\begin{aligned} \hat{H}_{\text{F}} &= \frac{1}{2} \int d^3x \left[ \frac{\hat{\mathbf{\Pi}}^2}{\epsilon_0\epsilon(\mathbf{x})} + \frac{\hat{\mathbf{B}}^2}{\mu_0} \right] \\ &= \sum_{\mu} \hbar\omega_{\mu} \hat{a}_{\mu}^{\dagger} \hat{a}_{\mu}, \end{aligned} \quad (\text{D21})$$

where in the second line we have used the expansions in Eq. (D16) and dropped the zero point energy, and

$$\hat{V}_{\text{Coul}} = - \int d^3x \frac{\hat{\mathbf{P}}^{\text{gC}} \cdot \hat{\mathbf{P}}^{\text{gC}}}{2\epsilon_0\epsilon(\mathbf{x})}. \quad (\text{D22})$$

Explicit expressions can be derived for  $\hat{\mathbf{P}}^{\text{gC}}$ , and thus  $\hat{V}_{\text{Coul}}$  using the following method. First, note that the longitudinal part is already fixed by Gauss's law:

$$\begin{aligned} \hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x}) &= - \int d^3x' \mathbf{K}_{\parallel}(\mathbf{x}, \mathbf{x}') \hat{\rho}_A(\mathbf{x}') \\ &= \nabla^x \int d^3x' \frac{\hat{\rho}_A(\mathbf{x}')}{4\pi|\mathbf{x} - \mathbf{x}'|}. \end{aligned} \quad (\text{D23})$$

Now we can introduce the auxiliary function  $\hat{\mathbf{N}}(\mathbf{x}) = (\hat{\mathbf{P}}^{\text{gC}} - \hat{\mathbf{P}}_{\parallel}^{\text{gC}})/\epsilon(\mathbf{x})$ , which is generalized transverse. Thus, this function can be expanded in terms of the modal functions  $\hat{\mathbf{N}}(\mathbf{x}) = \sum_{\mu} \hat{n}_{\mu} \mathbf{h}_{\mu}(\mathbf{x})$ , where  $\hat{n}_{\mu}$  are operator-valued expansion coefficients. These can be found as

$$\begin{aligned} \hat{n}_{\mu} &= \int d^3x \mathbf{h}_{\mu}(\mathbf{x}) \cdot [\hat{\mathbf{P}}^{\text{gC}}(\mathbf{x}) - \hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x})] \\ &= - \int d^3x \mathbf{h}_{\mu}(\mathbf{x}) \cdot \hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x}), \end{aligned} \quad (\text{D24})$$

and thus we can find an explicit expression:

$$\begin{aligned} \hat{\mathbf{P}}^{\text{gC}}(\mathbf{x}) &= \hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x}) + \epsilon(\mathbf{x}) \hat{\mathbf{N}}(\mathbf{x}) \\ &= \hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x}) - \sum_{\mu} \epsilon(\mathbf{x}) \mathbf{h}_{\mu}(\mathbf{x}) \int d^3x' \hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x}') \cdot \mathbf{h}_{\mu}(\mathbf{x}) \\ &= \hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x}) - \epsilon(\mathbf{x}) \left[ \frac{\hat{\mathbf{P}}_{\parallel}^{\text{gC}}(\mathbf{x})}{\epsilon(\mathbf{x})} \right]_{\perp}. \end{aligned} \quad (\text{D25})$$

These results fully recover previous findings on quantization in the generalized Coulomb gauge [63–65,85], using a constrained quantization approach. Note that the electric field operator is, in the generalized Coulomb gauge,

$$\hat{\mathbf{E}}^{\text{gC}}(\mathbf{x}) = - \frac{1}{\epsilon_0\epsilon(\mathbf{x})} [\hat{\mathbf{\Pi}}(\mathbf{x}) + \hat{\mathbf{P}}^{\text{gC}}(\mathbf{x})]. \quad (\text{D26})$$

## 2. PZW transformation to generalized multipolar gauge

As in Sec. III we can transform from the generalized Coulomb gauge Hamiltonian  $\hat{H}^{\text{gC}}$  to the generalized multipolar Hamiltonian  $\hat{H}^{\text{gmp}}$  by the PZW operator:

$$\hat{H}^{\text{gmp}} = \hat{U}_{\text{PZW}}^{\dagger} \hat{H}^{\text{gC}} \hat{U}_{\text{PZW}}, \quad (\text{D27})$$

and the unitary operator takes the same form as the usual PZW transformation:

$$\hat{U}_{\text{PZW}} = \exp \left[ \frac{i}{\hbar} \int d^3x \hat{\mathbf{A}}(\mathbf{x}) \cdot \hat{\mathbf{Z}}_A(\mathbf{x}) \right]. \quad (\text{D28})$$

Note we have used the variable  $\mathbf{Z}_A(\mathbf{x}) = \sum_{\alpha} q_{\alpha} \hat{\mathbf{r}}_{\alpha} \int_0^1 ds \delta(\mathbf{x} - s\hat{\mathbf{r}}_{\alpha})$  from the main text to represent a quantity usually referred to as the polarization in this context, rather than  $\mathbf{P}$ , to avoid confusion with the gauge-dependent polarization defined in the main text and previous subsection, which is used in the canonical transformation from  $\hat{\mathbf{\Pi}}_A$  to  $\hat{\mathbf{\Pi}}$ .

Using the commutation relation (D15), we can find the transformation of the canonical momentum coordinate for the field as

$$\hat{U}_{\text{PZW}}^{\dagger} \hat{\mathbf{\Pi}}(\mathbf{x}) \hat{U}_{\text{PZW}} = \hat{\mathbf{\Pi}}(\mathbf{x}) + \int d^3x' \hat{\mathbf{Z}}_A(\mathbf{x}') \cdot \mathbf{Q}(\mathbf{x}', \mathbf{x}). \quad (\text{D29})$$

Note that the second term in Eq. (D29) is not the generalized transverse component of the polarization  $\hat{\mathbf{Z}}_A$ ; rather, the second term, when integrated against a vector function  $\mathbf{f}(\mathbf{x})$ , is an integral over the dot product of  $\hat{\mathbf{Z}}_A$  and the generalized transverse component of  $\mathbf{f}(\mathbf{x})$ .

Applying the PZW transformation to the Hamiltonian term involving  $\hat{\mathbf{\Pi}}$ , we thus obtain

$$\begin{aligned} & \hat{U}_{\text{PZW}}^{\dagger} \int d^3x \frac{\hat{\mathbf{\Pi}}^2}{2\epsilon_0\epsilon(\mathbf{x})} \hat{U}_{\text{PZW}} \\ &= \int d^3x \frac{\hat{\mathbf{\Pi}}^2}{2\epsilon_0\epsilon(\mathbf{x})} + \int d^3x \frac{\hat{\mathbf{\Pi}} \cdot \hat{\mathbf{Z}}_A}{\epsilon_0\epsilon(\mathbf{x})} + \int d^3x \frac{(\hat{\mathbf{Z}}_A - \hat{\mathbf{P}}^{\text{gC}})^2}{2\epsilon_0\epsilon(\mathbf{x})}. \end{aligned} \quad (\text{D30})$$

To arrive at the second term in Eq. (D30), we have used the fact that  $\hat{\mathbf{\Pi}}/\epsilon(\mathbf{x})$  is already generalized transverse. For the third term, we first note that we can use the identity

$$\epsilon(\mathbf{x}') Q_{ji}(\mathbf{x}', \mathbf{x}) = \epsilon(\mathbf{x}) Q_{ij}(\mathbf{x}, \mathbf{x}') \quad (\text{D31})$$

to write it as

$$\int d^3x \frac{\epsilon(\mathbf{x})}{2\epsilon_0} \left\{ \left[ \frac{\hat{\mathbf{Z}}_A(\mathbf{x})}{\epsilon(\mathbf{x})} \right]_{\perp}^{(\epsilon)} \right\}^2. \quad (\text{D32})$$

Then we use the decomposition  $\hat{\mathbf{Z}}_A = \hat{\mathbf{Z}}_{A,\perp} + \hat{\mathbf{Z}}_{A,\parallel}$ , and subsequently,

$$\left[ \frac{\hat{\mathbf{Z}}_A(\mathbf{x})}{\epsilon(\mathbf{x})} \right]_{\perp}^{(\epsilon)} = \frac{\hat{\mathbf{Z}}_{A,\perp}(\mathbf{x})}{\epsilon(\mathbf{x})} + \left[ \frac{\hat{\mathbf{Z}}_{A,\parallel}(\mathbf{x})}{\epsilon(\mathbf{x})} \right]_{\perp}^{(\epsilon)}. \quad (\text{D33})$$

Finally, we use Eq. (D25), and note that  $\hat{\mathbf{Z}}_{A,\parallel} = \hat{\mathbf{P}}_{\parallel}^{\text{gC}}$ . The entire generalized multipolar gauge Hamiltonian is then, neglecting the magnetic terms as in the main text,

$$\hat{H}^{\text{gmp}} = \sum_{\alpha} \frac{\hat{\mathbf{p}}_{\alpha}^2}{2m_{\alpha}} + \hat{H}_{\text{F}} + \int d^3x \frac{\hat{\mathbf{\Pi}} \cdot \hat{\mathbf{Z}}_A}{\epsilon_0\epsilon(\mathbf{x})} + \int d^3x \frac{\hat{\mathbf{Z}}_A^2}{2\epsilon_0\epsilon(\mathbf{x})}, \quad (\text{D34})$$

again in agreement with Refs. [63–65].

The electric field operator in the generalized multipolar gauge is, using Eqs. (D26) and (D29),

$$\begin{aligned} \hat{\mathbf{E}}^{\text{gmp}}(\mathbf{x}) &= \hat{\mathbf{E}}^{\text{gC}}(\mathbf{x}) - \int d^3x' \frac{\hat{\mathbf{Z}}_A(\mathbf{x}') \cdot \mathbf{Q}(\mathbf{x}', \mathbf{x})}{\epsilon_0\epsilon(\mathbf{x})} \\ &= \hat{\mathbf{E}}^{\text{gC}}(\mathbf{x}) - \left[ \frac{\hat{\mathbf{Z}}_A(\mathbf{x})}{\epsilon_0\epsilon(\mathbf{x})} \right]_{\perp}^{(\epsilon)} \end{aligned}$$

$$\begin{aligned} &= \hat{\mathbf{E}}^{\text{gC}}(\mathbf{x}) - \frac{1}{\epsilon_0\epsilon(\mathbf{x})} [\hat{\mathbf{Z}}_{A,\perp}(\mathbf{x}) - \hat{\mathbf{P}}_{\perp}^{\text{gC}}(\mathbf{x})] \\ &= [\hat{\mathbf{E}}_{\text{F}}(\mathbf{x})]_{\perp} - \frac{\hat{\mathbf{Z}}_A(\mathbf{x})}{\epsilon_0\epsilon(\mathbf{x})}, \end{aligned} \quad (\text{D35})$$

where in the second line we have used Eqs. (D11) and (D31), in the third line we have used Eqs. (D25) and (D33), and in the final line we have used defined  $[\hat{\mathbf{E}}_{\text{F}}(\mathbf{x})]_{\perp} = -\hat{\mathbf{\Pi}}(\mathbf{x})/[\epsilon_0\epsilon(\mathbf{x})]$  as the part of the transverse electric field operator that can be expanded in terms of bosonic operators, and Eq. (D26).

### 3. Material and mode truncation

In this subsection, we show how to introduce material and mode truncation in a manner which preserves gauge invariance. The procedure is essentially the same as in Sec. IV, but considering only two gauges (generalized Coulomb and generalized multipolar).

First, note that we can write the generalized Coulomb gauge Hamiltonian in the following form:

$$\hat{H}^{\text{gC}} = \hat{H}_{\text{F}} + \hat{U}_{\text{PZW}} \hat{H}_0 \hat{U}_{\text{PZW}}^{\dagger}, \quad (\text{D36})$$

where  $\hat{H}_0 = \sum_{\alpha} \frac{\hat{\mathbf{p}}_{\alpha}^2}{2m_{\alpha}} + \hat{V}_{\text{Coul}}$ , and again we have neglected the magnetic terms. From this form, it is easy to see that we can write the generalized multipolar Hamiltonian as

$$\hat{H}^{\text{gmp}} = \hat{U}_{\text{PZW}}^{\dagger} \hat{H}_{\text{F}} \hat{U}_{\text{PZW}} + \hat{H}_0. \quad (\text{D37})$$

As in Sec. IV B, we can introduce gauge-invariant interactions within the truncated material space by assuming  $\hat{H}_0$  to be represented by a discrete energy basis of a few material eigenstates, found in the absence of coupling with the medium. We then truncate the material particle position degrees of freedom in the operator  $\hat{U}_{\text{PZW}}$  to obtain the correct materially truncated Hamiltonian in either gauge. Note that this procedure, as pointed out in Ref. [65], neglects the influence of local variations in  $\epsilon(\mathbf{x})$  on the energy structure of the truncated material system. Assuming a material system of molecular or atomic scales, this variation can sometimes reduce the free-space Coulomb interaction between material particles by approximately a factor of  $1/\epsilon(\mathbf{x}_0)$ , at the location of the system of particles  $\mathbf{x}_0$ , but a full treatment of this requires a model of dispersion, and is beyond the scope of this formalism [65]. We can partially circumvent this problem by instead assuming the truncated  $\hat{H}_0$  to consist of the medium-modified energy levels, and focus on the interactions with the transverse field  $\hat{\mathbf{\Pi}}$ .

To truncate the mode expansion, we can use the mode truncation operator defined in Eq. (62); this is equivalent to letting the sum in the field expansions (D16) run over a finite number of modes, as well as in  $\hat{H}_{\text{F}} = \sum_{\mu} \hbar\omega_{\mu} \hat{a}_{\mu}^{\dagger} \hat{a}_{\mu}$ . As with material truncation, we should apply this directly to the unitary operators which generate the minimal coupling Hamiltonian.

Applying both material and mode truncation, then we find  $\hat{U}_{\text{PZW}} \rightarrow \hat{\mathcal{U}}_{\text{PZW}}$ , such that

$$\hat{\mathcal{U}}_{\text{PZW}} = \exp \left[ \frac{i}{\hbar} \int d^3x \hat{\mathcal{A}}(\mathbf{x}) \cdot \hat{\mathbf{Z}}_A(\mathbf{x}) \right]. \quad (\text{D38})$$

Thus, we find the correctly truncated generalized Coulomb gauge Hamiltonian:

$$\hat{\mathcal{H}}^{\text{gC}} = \hat{\mathcal{H}}_{\text{F}} + \hat{\mathcal{U}}_{\text{PZW}}^{\dagger} \hat{\mathcal{H}}_0 \hat{\mathcal{U}}_{\text{PZW}}, \quad (\text{D39})$$

and the correctly truncated multipolar gauge Hamiltonian

$$\hat{\mathcal{H}}^{\text{gmp}} = \hat{\mathcal{U}}_{\text{PZW}}^{\dagger} \hat{\mathcal{H}}_{\text{gC}} \hat{\mathcal{U}}_{\text{PZW}} = \hat{\mathcal{U}}_{\text{PZW}}^{\dagger} \hat{\mathcal{H}}_{\text{F}} \hat{\mathcal{U}}_{\text{PZW}} + \hat{\mathcal{H}}_0, \quad (\text{D40})$$

and, explicitly

$$\begin{aligned} \hat{\mathcal{U}}_{\text{PZW}}^{\dagger} \hat{\mathcal{H}}_{\text{F}} \hat{\mathcal{U}}_{\text{PZW}} &= \hat{\mathcal{H}}_{\text{F}} - \int d^3x [\hat{\mathcal{E}}_{\text{F}}(\mathbf{x})]_{\perp} \cdot \hat{\mathcal{Z}}_{\text{A}}(\mathbf{x}) \\ &+ \sum_{\mu} \frac{[\int d^3x \hat{\mathcal{Z}}_{\text{A}}(\mathbf{x}) \cdot \mathbf{h}_{\mu}(\mathbf{x})]^2}{2\epsilon_0}, \end{aligned} \quad (\text{D41})$$

where  $[\hat{\mathcal{E}}_{\text{F}}(\mathbf{x})]_{\perp} = \hat{\mathcal{P}}_{\text{M}}[\hat{\mathbf{E}}_{\text{F}}(\mathbf{x})]_{\perp} \hat{\mathcal{P}}_{\text{M}}$ . The difference with naive, direct mode truncation of the total Hamiltonian is in the third term (the polarization-squared term), as is clear from comparison with Eq. (D34).

- 
- [1] T. Gruner and D.-G. Welsch, Green-function approach to the radiation-field quantization for homogeneous and inhomogeneous Kramers-Kronig dielectrics, *Phys. Rev. A* **53**, 1818 (1996).
- [2] G. S. Agarwal, Quantum electrodynamics in the presence of dielectrics and conductors. I. Electromagnetic-field response functions and black-body fluctuations in finite geometries, *Phys. Rev. A* **11**, 230 (1975).
- [3] B. Huttner and S. M. Barnett, Quantization of the electromagnetic field in dielectrics, *Phys. Rev. A* **46**, 4306 (1992).
- [4] H. T. Dung, L. Knöll, and D.-G. Welsch, Three-dimensional quantization of the electromagnetic field in dispersive and absorbing inhomogeneous dielectrics, *Phys. Rev. A* **57**, 3931 (1998).
- [5] S. Scheel, L. Knöll, and D.-G. Welsch, QED commutation relations for inhomogeneous Kramers-Kronig dielectrics, *Phys. Rev. A* **58**, 700 (1998).
- [6] L. G. Suttorp and M. Wubs, Field quantization in inhomogeneous absorptive dielectrics, *Phys. Rev. A* **70**, 013816 (2004).
- [7] T. G. Philbin, Canonical quantization of macroscopic electromagnetism, *New J. Phys.* **12**, 123008 (2010).
- [8] G. S. Agarwal, Quantum electrodynamics in the presence of dielectrics and conductors. IV. General theory for spontaneous emission in finite geometries, *Phys. Rev. A* **12**, 1475 (1975).
- [9] P. T. Kristensen and S. Hughes, Modes and mode volumes of leaky optical cavities and plasmonic nanoresonators, *ACS Photonics* **1**, 2 (2014).
- [10] P. T. Kristensen, K. Herrmann, F. Intravaia, and K. Busch, Modeling electromagnetic resonators using quasinormal modes, *Adv. Opt. Photon.* **12**, 612 (2020).
- [11] L. G. Suttorp and A. J. van Wonderen, Fano diagonalization of a polariton model for an inhomogeneous absorptive dielectric, *Europhys. Lett.* **67**, 766 (2004).
- [12] S. Weinberg and S. Weinberg, *The Quantum Theory of Fields* (Cambridge University Press, Cambridge, 1995), Vol. 1.
- [13] H. T. Dung, L. Knöll, and D.-G. Welsch, Spontaneous decay in the presence of dispersing and absorbing bodies: General theory and application to a spherical cavity, *Phys. Rev. A* **62**, 053804 (2000).
- [14] M. Fleischhauer, Spontaneous emission and level shifts in absorbing disordered dielectrics and dense atomic gases: A Green's-function approach, *Phys. Rev. A* **60**, 2534 (1999).
- [15] P. Yao, C. Van Vlack, A. Reza, M. Patterson, M. M. Dignam, and S. Hughes, Ultrahigh Purcell factors and Lamb shifts in slow-light metamaterial waveguides, *Phys. Rev. B* **80**, 195106 (2009).
- [16] H. T. Dung, L. Knöll, and D.-G. Welsch, Resonant dipole-dipole interaction in the presence of dispersing and absorbing surroundings, *Phys. Rev. A* **66**, 063810 (2002).
- [17] P. T. Kristensen, J. Mørk, P. Lodahl, and S. Hughes, Decay dynamics of radiatively coupled quantum dots in photonic crystal slabs, *Phys. Rev. B* **83**, 075305 (2011).
- [18] S. Y. Buhmann, L. Knöll, D.-G. Welsch, and H. T. Dung, Casimir-Polder forces: A nonperturbative approach, *Phys. Rev. A* **70**, 052117 (2004).
- [19] M. Kamandar Dezfouli and S. Hughes, Quantum optics model of surface-enhanced Raman spectroscopy for arbitrarily shaped plasmonic resonators, *ACS Photonics* **4**, 1245 (2017).
- [20] A. Frisk Kockum, A. Miranowicz, S. De Liberato, S. Savasta, and F. Nori, Ultrastrong coupling between light and matter, *Nat. Rev. Phys.* **1**, 19 (2019).
- [21] P. Forn-Díaz, L. Lamata, E. Rico, J. Kono, and E. Solano, Ultrastrong coupling regimes of light-matter interaction, *Rev. Mod. Phys.* **91**, 025005 (2019).
- [22] D. De Bernardis, T. Jaako, and P. Rabl, Cavity quantum electrodynamics in the nonperturbative regime, *Phys. Rev. A* **97**, 043820 (2018).
- [23] D. De Bernardis, P. Pilar, T. Jaako, S. De Liberato, and P. Rabl, Breakdown of gauge invariance in ultrastrong-coupling cavity QED, *Phys. Rev. A* **98**, 053819 (2018).
- [24] M. A. D. Taylor, A. Mandal, P. Huo, and P. Huo, Resolving ambiguities of the mode truncation in cavity quantum electrodynamics, *Opt. Lett.* **47**, 1446 (2022).
- [25] A. F. Starace, Length and velocity formulas in approximate oscillator-strength calculations, *Phys. Rev. A* **3**, 1242 (1971).
- [26] J. Keeling, Coulomb interactions, gauge invariance, and phase transitions of the Dicke model, *J. Phys.: Condens. Matter* **19**, 295213 (2007).
- [27] O. Di Stefano, A. Settineri, V. Macrì, L. Garziano, R. Stassi, S. Savasta, and F. Nori, Resolution of gauge ambiguities in ultrastrong-coupling cavity quantum electrodynamics, *Nat. Phys.* **15**, 803 (2019).
- [28] A. Stokes and A. Nazir, Gauge ambiguities imply Jaynes-Cummings physics remains valid in ultrastrong coupling QED, *Nat. Commun.* **10**, 499 (2019).
- [29] M. A. D. Taylor, A. Mandal, W. Zhou, and P. Huo, Resolution of Gauge Ambiguities in Molecular Cavity Quantum Electrodynamics, *Phys. Rev. Lett.* **125**, 123602 (2020).

- [30] K. G. Wilson, Confinement of quarks, *Phys. Rev. D* **10**, 2445 (1974).
- [31] S. Savasta, O. Di Stefano, A. Settineri, D. Zueco, S. Hughes, and F. Nori, Gauge principle and gauge invariance in two-level systems, *Phys. Rev. A* **103**, 053703 (2021).
- [32] Y.-C. Han and L. B. Madsen, Comparison between length and velocity gauges in quantum simulations of high-order harmonic generation, *Phys. Rev. A* **81**, 063430 (2010).
- [33] J. Li, D. Golez, G. Mazza, A. J. Millis, A. Georges, and M. Eckstein, Electromagnetic coupling in tight-binding models for strongly correlated light and matter, *Phys. Rev. B* **101**, 205140 (2020).
- [34] B. J. Dalton, S. M. Barnett, and P. L. Knight, Quasi mode theory of macroscopic canonical quantization in quantum optics and cavity quantum electrodynamics, *J. Mod. Opt.* **46**, 1315 (1999).
- [35] D. P. Fussell and M. M. Dignam, Quasimode-projection approach to quantum-dot-photon interactions in photonic-crystal-slab coupled-cavity systems, *Phys. Rev. A* **77**, 053805 (2008).
- [36] C. Vendromin and M. M. Dignam, Nonlinear optical generation of entangled squeezed states in lossy nonorthogonal quasimodes: An analytic solution, [arXiv:2210.06521](https://arxiv.org/abs/2210.06521).
- [37] E. S. C. Ching, P. T. Leung, A. Maassen van den Brink, W. M. Suen, S. S. Tong, and K. Young, Quasinormal-mode expansion for waves in open systems, *Rev. Mod. Phys.* **70**, 1545 (1998).
- [38] P. Lalanne, W. Yan, K. Vynck, C. Sauvan, and J.-P. Hugonin, Light interaction with photonic and plasmonic resonances, *Laser Photonics Rev.* **12**, 1700113 (2018).
- [39] S. Franke, S. Hughes, M. K. Dezfouli, P. T. Kristensen, K. Busch, A. Knorr, and M. Richter, Quantization of Quasinormal Modes for Open Cavities and Plasmonic Cavity Quantum Electrodynamics, *Phys. Rev. Lett.* **122**, 213901 (2019).
- [40] S. Hughes, S. Franke, C. Gustin, M. Kamandar Dezfouli, A. Knorr, and M. Richter, Theory and limits of on-demand single-photon sources using plasmonic resonators: A quantized quasinormal mode approach, *ACS Photonics* **6**, 2168 (2019).
- [41] S. Franke, M. Richter, J. Ren, A. Knorr, and S. Hughes, Quantized quasinormal-mode description of nonlinear cavity-QED effects from coupled resonators with a Fano-like resonance, *Phys. Rev. Res.* **2**, 033456 (2020).
- [42] S. Franke, J. Ren, and S. Hughes, Quantized quasinormal-mode theory of coupled lossy and amplifying resonators, *Phys. Rev. A* **105**, 023702 (2022).
- [43] J. Ren, S. Franke, and S. Hughes, Quasinormal Modes, Local Density of States, and Classical Purcell Factors for Coupled Loss-Gain Resonators, *Phys. Rev. X* **11**, 041020 (2021).
- [44] D. Lentrod and J. Evers, *Ab Initio* Few-Mode Theory for Quantum Potential Scattering Problems, *Phys. Rev. X* **10**, 011008 (2020).
- [45] M. Babiker and L. Rodney, Derivation of the Power-Zienau-Woolley Hamiltonian in quantum electrodynamics by gauge transformation, *Proc. R. Soc. Lond. A* **385**, 439 (1983).
- [46] R. G. Woolley, Power-Zienau-Woolley representations of non-relativistic QED for atoms and molecules, *Phys. Rev. Res.* **2**, 013206 (2020).
- [47] R. G. Woolley, Charged particles, gauge invariance, and molecular electrodynamics, *Int. J. Quantum Chem.* **74**, 531 (1999).
- [48] E. Rousseau and D. Felbacq, The quantum-optics Hamiltonian in the multipolar gauge, *Sci. Rep.* **7**, 11115 (2017).
- [49] A. Vukics, G. Kónya, and P. Domokos, The gauge-invariant Lagrangian, the Power-Zienau-Woolley picture, and the choices of field momenta in nonrelativistic quantum electrodynamics, *Sci. Rep.* **11**, 16337 (2021).
- [50] D. L. Andrews, G. A. Jones, A. Salam, and R. G. Woolley, Perspective: Quantum Hamiltonians for optical interactions, *J. Chem. Phys.* **148**, 040901 (2018).
- [51] A. Stokes and A. Nazir, Identification of Poincaré-gauge and multipolar nonrelativistic theories of QED, *Phys. Rev. A* **104**, 032227 (2021).
- [52] C. W. Gardiner and M. J. Collett, Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation, *Phys. Rev. A* **31**, 3761 (1985).
- [53] W. Salmon, C. Gustin, A. Settineri, O. Di Stefano, D. Zueco, S. Savasta, F. Nori, and S. Hughes, Gauge-independent emission spectra and quantum correlations in the ultrastrong coupling regime of open system cavity-QED, *Nanophotonics* **11**, 1573 (2022).
- [54] I. Medina, F. J. García-Vidal, A. I. Fernández-Domínguez, and J. Feist, Few-Mode Field Quantization of Arbitrary Electromagnetic Spectral Densities, *Phys. Rev. Lett.* **126**, 093601 (2021).
- [55] M. Sánchez-Barquilla, F. J. García-Vidal, A. I. Fernández-Domínguez, and J. Feist, Few-mode field quantization for multiple emitters, *Nanophotonics* **11**, 4363 (2022).
- [56] J. del Pino, F. A. Y. N. Schröder, A. W. Chin, J. Feist, and F. J. García-Vidal, Tensor Network Simulation of Non-Markovian Dynamics in Organic Polaritons, *Phys. Rev. Lett.* **121**, 227401 (2018).
- [57] S. Arranz Regidor, G. Crowder, H. Carmichael, and S. Hughes, Modeling quantum light-matter interactions in waveguide QED with retardation, nonlinear interactions, and a time-delayed feedback: Matrix product states versus a space-discretized waveguide model, *Phys. Rev. Res.* **3**, 023030 (2021).
- [58] A. Mercurio, V. Macrì, C. Gustin, S. Hughes, S. Savasta, and F. Nori, Regimes of cavity QED under incoherent excitation: From weak to deep strong coupling, *Phys. Rev. Res.* **4**, 023048 (2022).
- [59] A. Settineri, O. Di Stefano, D. Zueco, S. Hughes, S. Savasta, and F. Nori, Gauge freedom, quantum measurements, and time-dependent interactions in cavity QED, *Phys. Rev. Res.* **3**, 023079 (2021).
- [60] A. Stokes and A. Nazir, Ultrastrong time-dependent light-matter interactions are gauge relative, *Phys. Rev. Res.* **3**, 013116 (2021).
- [61] S. Savasta, O. Di Stefano, and F. Nori, Thomas-Reiche-Kuhn (TRK) sum rule for interacting photons, *Nanophotonics* **10**, 465 (2021).
- [62] A. Stokes and A. Nazir, Implications of gauge freedom for nonrelativistic quantum electrodynamics, *Rev. Mod. Phys.* **94**, 045003 (2022).
- [63] B. J. Dalton, E. S. Guerra, and P. L. Knight, Field quantization in dielectric media and the generalized multipolar hamiltonian, *Phys. Rev. A* **54**, 2292 (1996).
- [64] B. J. Dalton and M. Babiker, Macroscopic quantization in quantum optics and cavity quantum electrodynamics: Interatomic interactions, *Phys. Rev. A* **56**, 905 (1997).
- [65] M. Wubs, L. G. Suttorp, and A. Lagendijk, Multipole interaction between atoms and their photonic environment, *Phys. Rev. A* **68**, 013822 (2003).

- [66] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms: Introduction to Quantum Electrodynamics* (Wiley-VCH, Weinheim, 1997).
- [67] K. Sundermeyer, *Constrained Dynamics with Applications to Yang-Mills Theory, General Relativity, Classical Spin, Dual String Model* (Springer, Berlin, Heidelberg, 1982).
- [68] D. M. Rouse, B. W. Lovett, E. M. Gauger, and N. Westerberg, Avoiding gauge ambiguities in cavity quantum electrodynamics, *Sci. Rep.* **11**, 4281 (2021).
- [69] S. Savasta and R. Girlanda, The particle-photon interaction in systems described by model Hamiltonians in second quantization, *Solid State Commun.* **96**, 517 (1995).
- [70] M. Graf and P. Vogl, Electromagnetic fields and dielectric response in empirical tight-binding theory, *Phys. Rev. B* **51**, 4940 (1995).
- [71] T. Onodera, E. Ng, C. Gustin, N. Lörch, A. Yamamura, R. Hamerly, P. L. McMahon, A. Marandi, and H. Mabuchi, Non-linear quantum behavior of ultrashort-pulse optical parametric oscillators, *Phys. Rev. A* **105**, 033508 (2022).
- [72] S. M. Dutra and G. Nienhuis, Quantized mode of a leaky cavity, *Phys. Rev. A* **62**, 063805 (2000).
- [73] C. Viviescas and G. Hackenbroich, Field quantization for open optical cavities, *Phys. Rev. A* **67**, 013805 (2003).
- [74] D. M. Gitman and I. V. Tyutin, *Quantization of Fields with Constraints* (Springer, Berlin, 1990).
- [75] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1980).
- [76] L. Garziano, A. Settineri, O. Di Stefano, S. Savasta, and F. Nori, Gauge invariance of the Dicke and Hopfield models, *Phys. Rev. A* **102**, 023718 (2020).
- [77] C. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*, 3rd ed. (Springer, Berlin, 2004).
- [78] C. Sánchez Muñoz, F. Nori, and S. De Liberato, Resolution of superluminal signalling in non-perturbative cavity quantum electrodynamics, *Nat. Commun.* **9**, 1924 (2018).
- [79] P. A. M. Dirac, Generalized Hamiltonian dynamics, *Proc. R. Soc. London A* **246**, 326 (1958).
- [80] S. Scheel and S. Y. Buhmann, Macroscopic QED—Concepts and applications, *Acta Phys. Slov.* **58**, 675 (2008).
- [81] S. Franke, J. Ren, S. Hughes, and M. Richter, Fluctuation-dissipation theorem and fundamental photon commutation relations in lossy nanostructures using quasinormal modes, *Phys. Rev. Res.* **2**, 033332 (2020).
- [82] J. Ren, S. Franke, A. Knorr, M. Richter, and S. Hughes, Near-field to far-field transformations of optical quasinormal modes and efficient calculation of quantized quasinormal modes for open cavities and plasmonic resonators, *Phys. Rev. B* **101**, 205402 (2020).
- [83] R.-C. Ge, P. T. Kristensen, J. F. Young, and S. Hughes, Quasinormal mode approach to modelling light-emission and propagation in nanoplasmonics, *New J. Phys.* **16**, 113048 (2014).
- [84] R.-C. Ge, J. F. Young, and S. Hughes, Quasi-normal mode approach to the local-field problem in quantum optics, *Optica* **2**, 246 (2015).
- [85] R. J. Glauber and M. Lewenstein, Quantum optics of dielectric media, *Phys. Rev. A* **43**, 467 (1991).
- [86] J. R. de Lasson, P. T. Kristensen, J. Mørk, and N. Gregersen, Semianalytical quasi-normal mode theory for the local density of states in coupled photonic crystal cavity-waveguide structures, *Opt. Lett.* **40**, 5790 (2015).
- [87] J. Ren, S. Franke, and S. Hughes, Connecting classical and quantum mode theories for coupled lossy cavity resonators using quasinormal modes, *ACS Photonics* **9**, 138 (2022).