1 Introduction

In 2007, a report compiled for the Office of Science of the Department of Energy (DOE) identified five grand challenges in basic energy science, including (1) controlling material processes at the level of electrons, (2) designing and perfecting atom- and energy-efficient syntheses of new forms of matter with tailored properties, (3) understanding and controlling the remarkable properties of matter that emerge from complex correlations of atomic and electronic constituents, (4) mastering energy and information on the nanoscale to create new technologies with capabilities rivaling those of living things, and (5) characterizing and controlling matter away—especially far away—from equilibrium. One of the emerging techniques to address (some of) these challenges is light-matter interactions, which can be used to monitor, manipulate, and design materials’ properties through the complex interplay between electrons, photons, and phonons.2,3

In conventional chemistry, chemists modify the functionalities of molecules based on different functional groups. For centuries, such a vision has been widely used in chemical synthesis to modify molecules or molecular materials’ chemical and physical properties. However, in the current era of technology, previously “standard” chemistry, now in a complex electromagnetic (EM) environment, has emergent properties due to multiple new types of couplings between the light and matter degrees of freedom that can be used to control chemistry. Light-matter interactions have been instrumental in many branches of physics, chemistry, materials, and energy science. In most previous applications, the magnitude and scale of the light–matter interaction fall within the so-called weak coupling regime and can usually be treated at the lowest order in quantum electrodynamics via many-body perturbation theory. Such treatments are widely used in different types of spectroscopy techniques, optoelectronics, quantum sensing, quantum information, light harvesting, and beyond.

Alternatively, light-matter interactions open multiple new avenues for manipulating matter through novel emerging elementary excitations, including plasmons and polaritons. Plasmons (either surface plasmons or localized surface plasmons) are the collective oscillations of conduction electrons in nanostructures, which can be excited when the frequency of external light matches the plasmon resonant energies. Plasmon excitation results in significantly amplified absorption and scattering cross-sections. Moreover, plasmon excitations can overcome the diffraction limit and concentrate the incident light into a highly localized volume, enhancing the electromagnetic (EM) field (or photons) by several orders of magnitude in the near field. The subsequent plasmon decay process generates hot electrons or heat through electron–electron and electron–phonon scatterings at different length and time scales. Nevertheless, the resulting locally enhanced EM field, hot electrons, and heat can stimulate chemical reactions through various mechanisms. In fact, plasmon-mediated chemical reactions
reactions (namely plasmon chemistry) have become a promising strategy to drive chemical processes over the past decade.\textsuperscript{21} On the other hand, when molecules are collectively and resonantly coupled with plasmon excitations, a new quasiparticle (namely polaritons) can be formed in the strong coupling regime. The term “strong” coupling is relative, meaning that the light-matter coupling is large enough to compete with or overcome dissipation or dephasing (i.e., when the coherent energy exchange between a confined light mode and the quantum matter is faster than the decay and decoherence time scales of each part). Such strong coupling can be achieved when either a plasmonic mode is coupled with a few molecules\textsuperscript{23} or many molecules are collectively coupled to a single cavity mode.\textsuperscript{24} In the strong coupling regime, photons and electron/excitonic excitations in matter become equally important and are strongly coupled on an equally quantized footing. As a consequence, individual “free” particles no longer exist. Instead, the fundamental excitations of the light-matter interacting system are polaritons, which are hybrid light-matter excitations (superpositions of quantized light and matter) (Fig. 1)\textsuperscript{3} and possess both light and matter characteristics/topologies. Experiments have shown that matter properties can be modified with the formed polaritons, resulting in different photophysics and photochemistry.\textsuperscript{25} Since photon energies and light-matter coupling strengths are relatively tunable through cavity control, light–matter interaction in the strong coupling regime provides a fundamentally new way to manipulate matter properties for various desired applications, including lasing,\textsuperscript{26,27} electronics,\textsuperscript{28} long-range energy transfer,\textsuperscript{29–32} Bose-Einstein condensates,\textsuperscript{23,33–35} chemical reactions,\textsuperscript{36–37} internal conversion,\textsuperscript{38–40} singlet fission,\textsuperscript{61–63} manipulation of conical intersection,\textsuperscript{54–74} etc. Nevertheless, in polariton chemistry, light and matter cannot be treated as separate entities, as the strong coupling between them dresses each of them. Consequently, previously developed quantum chemistry methods for treating the electronic structure of matter alone become invalid. Brand new theoretical and modeling capabilities for describing polariton chemistry are required.

Despite the attractive applications of strong light–matter interaction, there are many open questions and fundamental problems about the mechanisms of physics and chemistry mediated by the strong light–matter interaction. The experimental progress should be complemented by theoretical advances that can offer atomistic insights into the problems that are not accessible by experiments alone. Progress in understanding the coupling between photons and elementary quasiparticles (plasmons, phonons, and excitons) in materials requires a generalized treatment of photons as one of the core DOFs in light–matter interactions. Unfortunately, the strong light–matter interaction is fundamentally a multiscale and multiphysics problem that involves multiple interactions between many DOFs and their complex interplay with environments across different spatial and time scales (Fig. 1). The traditional perturbation methods in the weak coupling regime are not applicable to the strong coupling regime, making it urgent to develop new theoretical and modeling techniques, especially multiscale methods, to understand and ultimately predict strong light–matter interaction-mediated physics and chemistry.

This perspective reviews recent theoretical efforts toward understanding the underlying mechanisms of plasmon and polariton chemistry due to the complex light-matter interactions and discusses our thoughts on future work. The article is structured as follows: Section 2 introduces the framework and basic mathematical structure of molecular quantum electrodynamics theory (QED), Section 3 explores the theoretical development in simulating plasmonic cavities and their strong interactions with molecules, Section 4 surveys recent progress in polariton chemistry in Fabry-Pérot-like cavities, and Section 5 concludes the discussion and provides a final perspective on future work in all realms of strong light–matter interaction.

2 Brief introduction to light-matter interactions and molecular QED theory

This section briefly introduces the quantum theories of light-matter interactions across different coupling regimes in the nonrelativistic limit.\textsuperscript{75} In principle, all light-matter interactions arise from the interplay of matter DOFs (degrees of freedom) (electrons or spins, nuclei) and an EM environment. Hence, a full \textit{ab initio} theory for light-matter interactions fundamentally requires the electronic structure theory of matter and principles of electrodynamics. Quantum electrodynamics (QED) is the
indispensable and most precise theory for describing the interactions of charged particles and the dynamics of the EM field in mutual interaction. Hence, we start with a brief introduction to the universal molecular QED theory, followed by its derivation of ab initio methods for plasmon and polariton chemistry (in different coupling regimes), respectively.

### 2.1 Molecular quantum electrodynamics theory

The minimally coupled Coulomb Hamiltonian governs the nonrelativistic dynamics of matter in an EM environment,

$$
\hat{H} = \sum_{i} \frac{\hbar}{2m_i} \left[ \hat{p}_i - z_i \mathbf{A}(\mathbf{r}_i) / c \right]^2 + \hat{V} + \hat{H}_{\text{EM}},
$$

(1)

with the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$. $\hat{p}_i = -i\hbar \nabla_i$ and $z_i$ are the momentum operator and charge of particle $i$, respectively. $N_e$ and $N_n$ are the number of electrons and nuclei in the systems. $\hat{V} = \hat{V}_{\text{ee}} + \hat{V}_{\text{en}} + \hat{V}_{\text{nn}} + \hat{V}_{\text{ext}}$ contains all of the Coulomb interactions $\left( \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|} \right)$ between the electronic and nuclear DOFs and another external potential. The Hamiltonian that describes the EM fields is

$$
\hat{H}_{\text{EM}} = \frac{\hbar^2}{2} \int d\mathbf{r} \left[ \mathbf{E}^2(\mathbf{r}) + \mathbf{c}^2 \mathbf{B}^2(\mathbf{r}) \right].
$$

(2)

It obeys Maxwell’s equations of motion and couples with the Schrödinger equation self-consistent through the vector potential $\mathbf{A}(\mathbf{r})$.

The general quantization of electromagnetic field involves three steps: (1) classical mode description of the electromagnetic field, (2) discretization of classical modes via standard mode decomposition techniques, and (3) quantization via the correspondence principle. After the quantization of the electromagnetic field, each field (including the vector potential $\mathbf{A}$) can be expressed as a sum over all possible modes

$$
\mathbf{A}(\mathbf{r}, t) = \sum_x u_x e^{i(\mathbf{K} \cdot \mathbf{r} - \omega_x t)} + \sum_x A_x^\dagger e^{i(\mathbf{K} \cdot \mathbf{r} + \omega_x t)}
$$

(3)

$$
\mathbf{E}(\mathbf{r}, t) = i \sum_x \omega_x u_x e^{i(\mathbf{K} \cdot \mathbf{r} - \omega_x t)} + A_x^\dagger e^{i(\mathbf{K} \cdot \mathbf{r} + \omega_x t)}
$$

(4)

$$
\mathbf{B}(\mathbf{r}, t) = i \sum_x (\nabla \times \mathbf{u}_x) e^{i(\mathbf{K} \cdot \mathbf{r} - \omega_x t)} + \frac{\omega_x}{c} A_x^\dagger e^{i(\mathbf{K} \cdot \mathbf{r} + \omega_x t)}
$$

(5)

where $\nabla \times \mathbf{u} = \mathbf{K}$ and $x \equiv \mathbf{x} \mathbf{K}$ denotes the photon mode with momentum $\mathbf{K}$ and polarization $x \in \{-1, 1\}$, $\mathbf{u}$ is the unit vector denoting the direction of the vector potential, and $A_x$ ($A_x^\dagger$) is the mode decomposition coefficients that create (annihilate) the $x$th radiation mode. Note that $\hat{\mathbf{E}} = -\frac{1}{c} \partial_t \mathbf{A}(\mathbf{r})$ and $\hat{\mathbf{B}} = \frac{1}{c} \nabla \times \mathbf{A}(\mathbf{r})$ in the Coulomb gauge (i.e., $\nabla \cdot \mathbf{A} = 0$). The photon modes in a given nanophotonic/nanoplasmonic structure can be readily calculated via standard mode decomposition techniques.

In the second quantization, the photonic Hamiltonian can then be written as

$$
\hat{H}_{\text{p}} = \frac{\hbar \omega_x}{2} (\hat{a}_x^\dagger \hat{a}_x + \text{h.c.}) = \frac{\hbar \omega_x}{2} (\hat{a}_x^\dagger \hat{a}_x + 1/2).
$$

with the equivalence,

$$
A_x \equiv \sqrt{\frac{\hbar}{2 k_0 V}} \hat{a}_x, \quad A_x^\dagger \equiv \sqrt{\frac{\hbar}{2 k_0 V}} \hat{a}_x^\dagger.
$$

(6)

Thus, the fields are quantized by the isomorphic association of a quantum mechanical harmonic oscillator to each radiation mode $x$. Introducing the canonical position and momentum operators as, $\hat{q}_x = \sqrt{\frac{\hbar}{2 k_0 V}} (\hat{a}_x^\dagger + \hat{a}_x)$, $\hat{p}_x = i \frac{\hbar}{2 k_0 V} (\hat{a}_x^\dagger - \hat{a}_x)$. The photonic Hamiltonian can now be rewritten as,

$$
\hat{H}_{\text{p}} = \frac{1}{2} \sum_x (\hat{p}_x^2 + \omega_x^2 \hat{q}_x^2).
$$

(7)

Then the full light-matter Hamiltonian in Coulomb gauge can now be fully constructed and written as,

$$
\hat{H}_c = \sum_{i} \frac{\hbar}{2m_i} \left[ \hat{p}_i - z_i \mathbf{A}(\mathbf{r}_i) / c \right]^2 + \hat{V} + \hat{H}_{\text{p}},
$$

(8)

which is often referred to as the minimal coupling Hamiltonian.

We now aim to apply a unitary transformation on eqn (8) to achieve an expression where the momenta of the molecular DOFs ($\hat{p}_i$) are decoupled from the vector potential ($\mathbf{A}(\mathbf{r})$). In other words, we will shift the light-matter coupling from momentum fluctuations into displacement fluctuations. This transformation is written as $\hat{U}(\mathbf{D}) = \exp \left[ -\frac{i}{\hbar} \mathbf{D} \cdot \mathbf{A} \right]$ where $\mathbf{D} = \sum_i \mathbf{r}_i \varepsilon_i - \sum_z e \mathbf{r}_z$ is the molecular dipole moment. This transformation $\hat{U}(\mathbf{D})\hat{H}_c\hat{U}(\mathbf{D})$ is nothing but a reduction in matter momentum such that $\hat{p}_i - z_i \mathbf{A} \rightarrow \hat{p}_i$ and a boost in photonic momentum by $\hat{p}_x \rightarrow \hat{p}_x + \sqrt{2 \omega / \hbar} \mathbf{D} \cdot \mathbf{A}_0$. Applying the transformation results in the QED Hamiltonian, referred to as the dipole gauge Hamiltonian,

$$
\hat{H} = \hat{T} + \hat{V} + \hat{H}_{\text{ep}}
$$

(9)

where

$$
\hat{H}_{\text{ep}} = \hat{U}^\dagger \hat{H}_p \hat{U} = \frac{1}{2} \sum_x \left[ (\hat{p}_x + \lambda_x \mathbf{D})^2 + \omega_x^2 (\hat{q}_x)^2 \right]
$$

$$
= \frac{1}{2} \sum_x \left[ \hat{p}_x^2 + \omega_x^2 (\hat{q}_x - \frac{\lambda_x}{\omega_x} \mathbf{D})^2 \right].
$$

(10)

The second line is reached via canonical transformation between coordinate and momentum operators, $\hat{p} \rightarrow -i \omega \hat{q} \rightarrow 1/\omega \hat{p}$. Here, $\lambda_x = \frac{1}{c} \varepsilon_x e \mathbf{u}_x$ and $g_x = \lambda_x \mathbf{D}$ defines the light-matter coupling strength, which depends on the volume of quantized photon $V$ in radiation mode $x$. The total light-matter Hamiltonian can now be re-partitioned after the unitary transformation as,

$$
\hat{H}_{\text{ME}} = \hat{H}_M + \hat{H}_p + \hat{H}_{\text{ep}} + \hat{H}_{\text{DSE}}
$$

$$
= \hat{H}_M + \sum_x \left[ \omega_x \left( \hat{a}_x^\dagger \hat{a}_x + \frac{1}{2} \right) + \sqrt{\frac{\omega_x}{2}} \mathbf{D} (\hat{a}_x^\dagger + \hat{a}_x) + \frac{1}{2} (\lambda_x \mathbf{D})^2 \right].
$$

(11)

(12)

Here, $\hat{H}_M = \hat{T}_M + \hat{V}_M + \hat{\mathbf{V}}$ is the bare molecular Hamiltonian which
includes all Coulomb interactions $\hat{V}$ between electrons and nuclei as well as the kinetic energy operators of both, $\hat{T}_e$ and $\hat{T}_n$, respectively. This Hamiltonian is often referred to as the Pauli-Fierz (PF) Hamiltonian.

Now it is clear that the light-matter coupling strength is determined by the two quantities: the molecular dipole strength and the quantized cavity volume. Hence, there are two general strategies to enter the strong coupling regime: (I) reduce the cavity volume and (II) increase the number of molecules (increasing the total dipole moment). Therefore, there are two major experimental nanocavity designs for strong light-matter coupling. The first is the nanophotonic cavity that leverages a large number of molecules to achieve strong coupling. The other is the nanoplasmmonic cavity that leverages a locally enhanced electric field confined in a small volume to enhance the coupling. This local field is effectively confined to the nearby surrounding of a spherical nanoparticle residing on a lattice of nanoparticles on the scale of nm$^3$ or even Å$^3$ (namely picocavities$^{23,80,81}$).

The derivation of eqn (10) assumes dipole approximation. However, the dipole approximation may fail in the ultra-confined nanoplasmmonic cavities where the EM fields are confined within nanometric volumes.$^{82,83}$ Consequently, the size of molecules becomes comparable to the cavity volumes, and the widely used dipole approximation breaks down. Position-dependent coupling strength that requires the spatial distribution of excitonic and photonic quantum states is found to be a key aspect in determining the dynamics in ultrasmall cavities both in the weak and strong coupling regimes.$^{82}$

2.2 Weak and strong coupling

The light-matter interactions in nanoplasmmonic environments can be split into weak and strong coupling regimes. The weak coupling regime is associated with the Purcell enhancement of spontaneous emission. This effect is particularly strong when the molecule is placed next to a metallic surface or nanostructure.$^{84}$ The plasmon can tune photophysics and photochemistry (i.e., plasmon chemistry) via various possible pathways in such a regime, leading to a promising approach for accelerating and manipulating chemical reactions, which will be the topics of Section 3. In this regime, the light–matter interaction can be treated semiclassically via the classical treatment of the electromagnetic field,

$$H_{EM} = \frac{\varepsilon_0}{2} \int \left[ E^2(r) + c^2 B^2(r) \right].$$

(13)

$$\hat{H}_{ep} = \int E(r) \cdot (\text{cr}).$$

(14)

where $E(r)$ and $B(r)$ are classical electromagnetic fields. Afterward, the conventional quantum chemistry method can be utilized and extended to describe the plasmon-induced chemical processes. But it should be noted that, even in the weak coupling regime, there is growing interest in exploring the quantum properties of plasmon for applications.$^{76,85-87}$ More details will be discussed in Section 3.

The other regime is the strong coupling regime, where the light–matter interaction cannot be treated perturbatively. The strong coupling is characterized by a reversible (coherent) exchange of energy (known as Rabi oscillation) between the matter and the cavity photon because the coupling is strong enough to compete with the dissipation. In this regime, polariton formation requires theoretical methods that treat the matter and photonic DOFs on equal footing and describe the multiple interactions between photons, electrons, and nuclei on different length and time scales (polariton chemistry). The theoretical advances towards understanding polariton chemistry and our perspective on future multiscale modeling of polariton chemistry are discussed in Section 4.

3 Plasmon chemistry

Although confined plasmonic modes result in localized and enhanced fields, plasmons usually suffer from strong dissipation, making it hard to enter the strong coupling regime in plasmonic systems. Nevertheless, even without strong coupling, plasmonics provides a unique setting for manipulating light via the confinement of EM (below the diffraction limit). Such extreme concentration of EM field$^{86}$ has led to a wide range of applications, such as plasmon-enhanced molecular spectroscopy$^{8,89-92}$, photovoltaics$^{93,94}$, nanophotonic lasers and amplifiers$^{95-98}$, quantum information$^{16}$, and many others.$^{18,99,100}$

Following the ultrafast plasmon excitation, nonradiative plasmon decay leads to the formation of energetic electron-hole pairs (namely hot-carriers)$^{101-103}$ which are highly non-thermal and can have considerably higher energies than those rising from thermal equilibrium. The hot electrons (HEs) (and their corresponding holes) redistribute their energies quickly as a result of electron–electron scattering$^{101,102}$ reaching a quasithermal equilibrium but with an effective high temperature. Further cooling of the hot electrons takes place via energy dissipation into the phonon modes of the nanoparticle, and the energy is ultimately dissipated to the surroundings via thermal conduction. Nevertheless, investigations in the last two decades have found that chemical reactions can be stimulated by localized EM fields (or photons), and electronic and/or thermal energies (that result from plasmon decay) via various pathways (more details in Section 3.4). This leads to an emerging field of plasmon chemistry that designs nanostructure-based surface plasmons as mediators to redistribute and convert photon energy in various time, space, and energy scales to drive chemical reactions.$^{20-23,92,106-116}$ One extraordinary feature of Plasmon chemistry in the realm of chemical reaction mechanisms is its blend of facets from thermochemistry, photochemistry, and photocatalysis.$^{92}$ Unlike traditional reactions that typically focus on singular mechanisms, plasmon chemistry showcases the interaction between different mechanisms in complex electronic and optical environments. At the same time, the mixture of different mechanisms makes it very challenging to comprehend the underlying principles of
plasmon chemistry, making it considerably more complicated than transition chemistry. This unique interplay often results in a diverse distribution of reactive zones on substrates.92 To fully understand plasmon chemistry’s nuances, an interdisciplinary approach is essential, which should consider multiscale processes, the current state of the field, and the need for advanced experimental methods.

3.1 Experimental endeavours in plasmon-mediated chemical reactions: the need for theoretical insights

Plasmon chemistry stands at the intersection of innovation and discovery for boosting chemical transformation. However, despite the spreading interest, inherent complexities make it still very challenging for efficiency optimization. By leveraging knowledge from established domains such as plasmon-enhanced spectroscopy, many experimental strategies to bolster plasmon chemistry efficiency can be extrapolated. In this section, we discuss several typical experimental investigations into plasmon chemistry across diverse applications and illustrate the essential role of theoretical/numerical modeling in understanding the underlying mechanisms for further improvement. For a more expansive disquisition on the plasmon-mediated chemical reactions for various applications, enthusiasts are redirected to several recent reviews.112,117–119

Artificial photosynthesis. One of the most enticing applications of plasmon chemistry echoes the wonders of photosynthesis.18,120–122 The goal is to convert solar energy into useful chemicals, a task intricate due to its kinetic and thermodynamic demands. However, in advancements in the realm of solar conversion, a groundbreaking study has showcased the potential of autonomous plasmonic catalysts capable of water splitting under visible light,120 as shown in Fig. 2a, present promising trajectories for renewable energy solutions. This paradigm shift focuses on deriving charge carriers from these plasmonic structures. The research highlights the introduction of a solar water-splitting device engineered from a gold nanorod array. Unlike conventional methods, this device leverages the power of “hot electrons” initiated by stimulating surface plasmons within the gold nanostructure. A standout feature is each nanorod’s autonomous operation, eliminating the need for external wiring and its impressive capability to produce H₂ molecules under sunlight. The device’s resilience and long-term stability further cement its potential. While the initial solar-to-hydrogen efficiency aligns with that of early semiconductor-based water splitters, it remains on the lower side for real-world applications. Nevertheless, the research suggests prospective structural enhancements, emphasizing the modification of the nanorod configuration, that could potentially escalate its efficiency. An important note for consideration, despite the current efficiency metrics, is the device’s remarkable longevity, which overtakes some of the top-performing semiconductor-based alternatives. This study marks a pivotal step towards the innovative utilization of plasmonic devices in solar conversion.

Fig. 2 Experimental examples: (a) structure and mechanism of operation of the autonomous plasmonic solar water splitter. Adapted with permission from ref. 120. Copyright 2013 Springer Nature Limited. (b) Light-driven CO₂ reduction with single atomic site antenna-reactor plasmonic photocatalysts. Adapted with permission from ref. 123. Copyright 2020 the author(s), under exclusive license to Springer Nature Limited. (c) Schematic diagram of the proposed mechanism for the oxidation of PTAP to DMAB. Hot electrons generated from the plasmon transfer to the adsorbed O₂ molecules, generating ³O₂ that participated in the PATP oxidation to DMAB. Adapted with permission from ref. 124. Copyright 2015 John Wiley and Sons. (d) Schematic of the three-electrode electrochemical system for distinguishing thermal from hot electron effects by utilizing photoelectrochemical characterization under the chopped light. Adapted from ref. 125 under Creative Commons CC BY.
CO₂ reduction. Given the increasing global CO₂ levels, plasmon chemistry offers a promising solution. Plasmonic structures have shown potential in boosting the photo-reduction of CO₂, which is crucial for environmental cleanup and advancing green chemistry. In a new development, researchers have presented a plasmonic photocatalyst that blends a copper (Cu) nanoparticle “antenna” with individual atomic ruthenium (Ru) sites (Fig. 2b). This design allows methane dry reforming to happen at room temperature, using the energy from light. Unlike traditional thermocatalytic methods, this photocatalyst works effectively with light at normal conditions and has notable stability and selectivity in its actions. This difference from standard thermally-driven reactions is due to the creation of hot carriers, which improve the rate of carbon-hydrogen activation on Ru sites and speed up hydrogen release. It should also be noted that, different from conventional plasmonic structure, the design of antenna-reactor systems utilize plasmonic metal (the antenna) to collect and concentrate visible light energy and transfer that energy to a catalytic metal (that is, the reactor) to drive a chemical reaction. Because the reactors were not in direct contact with the plasmonic nanoparticles, it was argued that the energy transfer resulting in increased reaction rates could only take place via a field effect where the field from the plasmonic metal was felt by the catalytic reactors resulting in the excitation of charge carriers in reactors.

Organic transformation. Beyond the catalysis of inorganic reactions, as previously addressed, plasmonic NPs have showcased commendable performance in catalyzing organic transformations. These hold potential applications, spanning from the synthesis of basic commodity molecules to intricate pharmaceutical compounds. This approach boasts benefits such as superior selectivity, accelerated reaction rates, and comparatively lenient reaction conditions. An illustrative instance is the oxidation of p-aminothiophenoul (PATP) or p-nitrothiophenoul (PNTP) to dimercaptaoazobenzene (DMAB). This reaction has been the focal point of myriad studies in recent years. The hypothesized microscopic mechanism posits that plasmon-excited electrons transfer from plasmonic nanostructures to O₂ molecules, resulting in the emergence of activated oxygen species, which, in turn, catalyze the formation of DMAB, as shown in Fig. 2(c). Such a mechanism found experimental validation through experiments toggling specific conditions. The obvious evidence of PATP’s transformation into DMAB, when air permeates the PATP and plasmonic nanostructured system, is corroborated by the ascendant Raman modes of DMAB over time. In contrast, in the absence of O₂ or electron transfer (attained by coating the plasmonic NPs with an insulator like SiO₂), merely the Raman signals of PATP are. In addition, it was found that the plasmon-induced hot holes can trigger the oxidation of PATP, even in the absence of oxygen. Notably, despite the absence of a tangible PATP transformation on gold nanoparticles due to inefficient charge transfer from the plasmonic nanostructure directly to the adsorbates, PATP oxidation has been witnessed either in the presence or absence of O₂, specifically with Au nanostructures fortified with a TiO₂ shell. These reactions outperform their counterparts in sheer gold nanoparticle systems in terms of efficiency. Fascinatingly, it was unveiled that achieving the selective oxidation of PNTP to either DMAB or PNTP is possible, depending on whether the plasmonic NPs are reinforced by TiO₂ and the presence of UV-illumination. Under UV exposure on the bare plasmonic NPs, PATP’s oxidation to DMAB is observed. In contrast, there is no catalytic effect with the TiO₂ support. However, after applying UV light exposure, a transition to PNTP is detected with the TiO₂ support. For an expansive disquisition on the plasmon-mediated organic transformation, enthusiasts are redirected to a contemporary review in ref. 117.

Experimental limitations. Though there are many demonstrated applications of plasmon chemistry as a promising approach to accelerate or manipulate chemical processes, the underlying mechanisms are still unclear. The advances in experimental techniques have played a pivotal role in elucidating plasmon chemistry mechanisms. Techniques, such as transient absorption spectroscopy, offer insights into the domain of plasmonic hot-carriers. Nonetheless, the vast expanse of these reactions remains to be traversed for a holistic comprehension. Traversing the landscape of plasmon chemistry requires a comprehensive understanding of multiple coherent and dissipative processes, including the enhancement of electromagnetic near-fields, local heating effects, and charge-carrier excitation/transfer. These processes occur across varied scales, from femtoseconds to nanoseconds. Since all these effects can drive chemical processes simultaneously, the individual contributions of these elements to overarching reactions are challenging to discern. On the other hand, decoupling individual contributions is vital for designing better plasmon catalysts that synchronize the orchestration of these effects to improve efficiency. In particular, distinguishing the thermal effect from the hot electron effect has become an important topic and is the subject of many debates. Despite the complexity of distinguishing these effects, there are growing efforts to solve these key questions via various designs. One of these efforts utilized the unique photoelectrochemical behavior of a plasmonic Au nanoelectrode array as shown in Fig. 2d. The plasmonic photocurrent can be intensified at negative and positive potentials, with its direction determined by the selected potential. This allows the electrode to favor either reduction or oxidation reactions. The photocurrent is composed of a fast photoelectronic component and a slower photothermal component. The photoelectronic current aligns with the plasmon absorption spectrum and intensifies with light intensity, whereas the photothermal current displays a linear relationship only within specific light intensities. Consequently, this proposed design can differentiate between the photoelectronic and photothermal effects. However, the photoelectronic response measured in ref. 123 happens at a time scale several orders of magnitude larger than the hot-electron lifetime and thermal conduction time scales. More advanced time-resolved techniques are required to provide more comprehensive evidence.
Overall, the distinctiveness of plasmon chemistry is rooted in the interplay between molecules, incident photons, and plasmonic nanostructures. Though many recent breakthroughs have been made to demonstrate plasmon chemistry's application in various areas, further improvements require a deep microscopic understanding of the synergy of multiple effects at the electron level. Such understanding will play an indispensable role in the strategic design and fabrication of plasmonic nanostructures, optimization of surface or interface intermediaries, and their harmonized integration. To elucidate the intricate mechanism that governs plasmon chemistry, comprehensive investigations encompassing various time, spatial, and energy scales are essential, which cannot be achieved by experiments alone. It's crucial that advanced experimental methods, boasting high space-time resolutions, are synergized with microscopic theoretical modeling. Specifically, theoretical models capable of addressing all potential mechanisms on equal footing provide invaluable perspectives into quandaries that remain elusive to experimental endeavors alone. Intuitively, the convenience of modeling allows for the toggling of individual effects to scrutinize their ramifications on holistic reactions. The real theoretical challenge, however, lies in considering these multifaceted effects equitably, spanning varied time and length scales. In the following sections, we will review our theoretical efforts toward simulating all conceivable mechanisms on an equal footing, with an aspiration to understand plasmon chemistry at the electron level.

3.2 Ab initio method for quantum plasmonics

In most scenarios, the nonplasmonic properties are obtained by solving Maxwell’s equations with proper dielectric function and boundary conditions. However, Maxwell’s equations fail to describe the quantum effects (such as the tunneling in charge transfer mode and quantum confinement85), and quantum theory for plasmon (quantum plasmonics) is required in nanoscale.

From a quantum mechanical point of view, a plasmon is nothing but a specific collective excitation. Hence, the quantum chemistry method for excited states can be used to predict the optical properties of plasmonic nanostructures. The many-body perturbation theory, time-dependent density functional theory (TDDFT), or time-dependent density functional tight-binding (TDDFTB) are methods of choice to compute the plasmon excitations.141–149 Without losing generality, here we use TDDFT theory as an example to demonstrate the computation of plasmon excitations.

TDDFT is the formal extension of the Hohenberg–Kohn–Sham density functional theory (DFT). Within the DFT, The Hohenberg–Kohn (HK) theorem150 states that the ground-state electron density unambiguously defines the many-electron ground state for an N-electron system under the influence of an external potential. And there exists an energy functional that guarantees the variational principle can reach the ground state energy, though the exact functional is an unsolved problem,

\[ E[\rho] = T + V_H(\mathbf{r}) + E_{xc}[\rho] + \int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}. \]  

(15)

Where \( V_H(\mathbf{r}) = \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \) is the Hartree potential, \( E_{xc} \) is the exchange–correlation functional, which contributes <10% to the total energy, but describes the most critical correlation effects.151 But DFT is a ground-state theory. Light-matter interactions usually result in many elementary excitations within the matter. Among various many-body methods, TDDFT (the formal extension of DFT theory) is the preferred tool to evaluate the excited states and optical properties in extended molecular or condensed matter systems.

In the linear response regime, the induced electron density due to the light–matter interaction \( \delta \rho(\mathbf{r}) \) is given by

\[ \delta \rho(\mathbf{r}) = \int \chi(\mathbf{r}, \mathbf{r}') \delta V_{ext}(\mathbf{r}')d\mathbf{r}'. \]  

(16)

where \( \chi(\mathbf{r}, \mathbf{r}') \) is the polarizability describing the density response of the many-body ground state with respect to an external perturbation. In addition, according to Runge-Gross theorem,152 \( \delta \rho(\mathbf{r}) \) is also the induced density of the KS system, but due to a perturbation

\[ \delta \rho(\mathbf{r}) = \int \delta V_{Ks}(\mathbf{r}, \mathbf{r}')d\mathbf{r}'. \]  

(17)

\( \delta V_{Ks} \) here is the linear response of the KS electrons, which can be trivially evaluated from the KS orbitals and energies.

\[ \delta V_{Ks}(\mathbf{r}) = \delta V_{ex}(\mathbf{r}) + \sum_{\alpha} \delta \rho_{\alpha}(\mathbf{r}) + \delta V_{xc}(\mathbf{r}). \]  

(18)

The induced XC potential is given by \( \delta V_{xc}(\mathbf{r}) = \int f_{xc}(\mathbf{r}, \mathbf{r}')\delta \rho(\mathbf{r}')d\mathbf{r}' \), where \( f_{xc} \) is the dynamical XC kernel

\[ f_{xc}(\mathbf{r}, \mathbf{r}') = \frac{\delta V_{xc}(\mathbf{r})}{\delta \rho(\mathbf{r}')}. \]  

(19)

Hence, a linear equation for the induced density can be derived from eqn (16)–(19),

\[ [1 - \chi_{Ks}(\omega)f_{xc}(\omega)]\chi(\omega)\delta V_{ex}(\omega) = \chi_{Ks}(\omega)\delta V_{ex}(\omega). \]  

(20)

Where \( f_{xc}(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + f_{xc}(\mathbf{r}, \mathbf{r}') \). Casting eqn (20) into the matrix of coupled KS single excitations, it is possible to calculate the excitation energies \( \omega_i \) of the system (the poles of the response function) and transition densities (and oscillator strengths). Consequently, the response function \( \chi(\omega) \) can be rewritten as

\[ \chi(\mathbf{r}, \mathbf{r}', \omega) = 2\sum_{\alpha} \rho_{\alpha}(\mathbf{r})\rho_{\alpha}(\mathbf{r}')\zeta_{\alpha}(\omega), \]  

(21)

where \( \zeta_{\alpha}(\omega) = \frac{1}{\omega - \omega_{\alpha} + i\eta} - \frac{1}{\omega + \omega_{\alpha} - i\eta} \) and \( \eta \) represents a positive infinitesimal. The factor of 2 comes from summation over spin indices. The excitation energies in the system are denoted as \( \omega_{\alpha} \), which are calculated from the Casida method.153 \( \rho_{\alpha}(\mathbf{r}) \) is the transition density, which can be expanded in terms of electronic transitions between
occupied state \( i \) to unoccupied states \( a \),

\[
\rho_{ia}(r) = \sum_{\alpha} X_{ia}^* \psi_i(r) \psi_a(r) \left( \frac{\omega_a - \omega_i}{\omega_a} \right)^{1/2},
\]

(22)

where \( \omega_i \) are the KS eigenvalues, and the corresponding molecular orbitals are \( \psi_i \). And \( X_{ia}^* \) are the Casida transition coefficient from the occupied \( i \) state to the unoccupied a state of the \( \alpha \) excitation. By examining the nature of \( X_{ia}^* \), it is possible to distinguish between normal excitation and plasmonic excitation. Plasmonic excitations generally are characterized by collective transitions from occupied to virtual orbitals.\(^{155,156} \)

### 3.3 Hot electron generation and relaxation

Once the plasmon excitation is obtained, the electron–plasmon interaction can be described by the Hamiltonian

\[
\hat{H}_{\text{int}} = \frac{e}{2m_e} \int dr \hat{\Psi}^\dagger V_{\text{eff}}(r) \hat{\Psi},
\]

(23)

where \( V_{\text{eff}}(r) \) is the effective potential induced by the excitation, which can be calculated from the transition density \( \rho_{ia}(r) \) by following Lundqvist’s approach.\(^{154,155} \) The polarizability of the NP upon external excitation is given by eqn (21). And the induced potential is given by,

\[
V_{\text{eff}}(r) = \varepsilon^{-1}(r,r) \delta V_{\text{exc}}(r),
\]

(24)

where the dielectric function \( \varepsilon^{-1}(r,r) \) is given by

\[
\varepsilon^{-1}(r,r) = \delta(r,r) + \int dr f_{\text{hxc}}(r,r) \chi_{\text{coh}}(r; r, \omega).
\]

(25)

Thus, within the second quantization, electron-excitation Hamiltonian \( \hat{H}_{\text{exc}} \) reads

\[
\hat{H}_{\text{exc}} = \sum_{ij} \left[ \mathcal{M}_{ij} c_i^\dagger c_j \Theta + \text{h.c.} \right],
\]

(26)

where \( \mathcal{M}_{ij} \) is the electron-excitation coupling strength, describing the scattering of quasiparticles from state \( i \) into state \( j \) via the emission or absorption of excitation in the state \( s \). These elements are given by (see ESI for details)

\[
\mathcal{M}_{ij} = \zeta_s(\omega) E \cdot \Gamma \left[ \langle \psi_i(r) | \hat{V}_{\text{coh}}^\dagger (r) | \psi_j(r) \rangle \right]
\]

(27)

where

\[
\Gamma = \sqrt{\frac{2m_e}{\hbar^2} \sum_{ia} X_{ia}^* \left( \frac{\omega_a - \omega_i}{\omega_a} \right)^{1/2}} \mathbf{d}_a,
\]

\[
|\Gamma|^2 = \sqrt{\frac{2m_e}{\hbar^2} \sum_{ia} X_{ia}^* \left( \frac{\omega_a - \omega_i}{\omega_a} \right)}
\]

\[
\mathbf{d}_a = \text{is the oscillator vector}
\]

and \( |\Gamma|^2 \) is the oscillator strength. \( \hat{V}_{\text{coh}}(r) = \int dr \frac{\rho_{ia}(r)}{r-r'} \) is the Hartree potential induced by the transition density. Detailed derivation can be found in the ESI. Above equation actually includes both plasmon excitation (\( \zeta_s(\omega) E \cdot \mathbf{f} \)), screening effect and hot electron–hole pairs generation (\( \langle \psi_i(r) | \hat{V}_{\text{coh}}^\dagger (r) | \psi_j(r) \rangle \)).

Neglecting the screening effect would result in a much larger electron-plasmon coupling matrix and shorter HC lifetime distribution.\(^{155} \) In contrast, our model connects the widely used semipirical model and the recently developed quantum model with the plasmon excitation and HC generation treated on equal footing. Eqn (26) is the general formalism that describes the coupling between electrons and photoexcitation. If the external field matches the plasmon energy, plasmon resonance will be excited, and eqn (26) reduces to the electron–plasmon coupling.

**HC generation.** The electron–plasmon coupling describes the HC generation following plasmon decay. After the electron–plasmon coupling matrix is obtained, the HC generation can be readily calculated from the Fermi golden rule,\(^{101,105} \)

\[
\Gamma_{i-a}^{\text{ex}} = \frac{4}{\hbar} \sum_{s} \left| \mathcal{M}_{ia} \right|^2 \frac{\gamma_{\text{ex}}}{(\omega - \omega_i + \omega_0)^2 + \gamma_{\text{ex}}^2} \delta(\omega - \omega_i).
\]

(28)

\( \gamma_{\text{ex}} \) refers to the linewidth of the excitation. \( \omega \) is the excitation energy. \( \delta(\omega - \omega_i) \) describes the generalized photoexcitation. Hence, eqn (28) describes the photoexcitation and HC generation. When the energy of the external field \( \omega \) matches the plasmon energy, eqn (28) describes the HC generation from plasmon decay. Otherwise, it describes the HC generation from regular excitation.

**HC relaxation.** After generation from plasmon decay, the HCs will undergo a relaxation process mainly due to the electron–electron and electron–phonon scatterings. The dynamics of the HCs are investigated by propagating the density matrix \( \rho(t) \). The equation of motion (EOM) of \( \rho \) is subject to the following quantum Liouville Von-Neumann equation,\(^{158} \)

\[
i \hbar \partial_t \rho(t) = i \hbar \partial_t \rho(t)|_{\text{coh}} + \mathcal{L}_{\text{ee}}[t] + \mathcal{L}_{\text{ph}}[t] + \mathcal{L}_{\text{L}}[t].
\]

(29)

The above equation’s first part on the right-hand side (RHS) describes the coherent evolution of the density matrix between different states. The second and third parts on the RHS of eqn (29) represent the dissipations induced by the electron–electron and electron–phonon scattering effects, respectively. Finally, the last term on the RHS of eqn (29) describes the extraction of HCs. In general, in the presence of electron–electron and electron–phonon interaction, the dissipation can be described by a Liouville equation by employing the many-body perturbation theory (MBPT),\(^{159} \)

\[
\mathcal{L}(t) = \mathcal{J}(t) - \mathcal{J}(t).
\]

Where the dissipation matrix \( \mathcal{J} \) can be written in terms of Green’s functions \( G \) and self-energies \( \Sigma \), i.e.,

\[
\mathcal{J}(t) = i \int dt' \Sigma^{\text{ex}}(t', t) G^{\text{ex}}(t', t) - \Sigma^{\text{coh}}(t, t) G^{\text{coh}}(t', t).
\]

(30)

Hence, the key point is to develop approximations to the self-energies \( \Sigma^{\text{ex}}(t', t) \) and efficient numerical methods to compute the \( \mathcal{J} \). The detailed formalisms of \( \mathcal{J} \) for electron–electron and electron–phonon scatterings can be found in ref. 101. But it should be noted that the approximation used for electron–electron scattering should conserve both particles and energy.

### 3.4 Plasmon-mediated chemical reactivities and theoretical challenges

However, the computations of plasmon excitation and distribution of hot electrons (from plasmon decay) are insufficient to give insights into plasmon-mediated chemistry because plasmon-induced photocatalysis is a complex dynamical
process that involves multiple interactions and mechanisms. To date, four major microscopic mechanisms have been proposed to explain how plasmonics facilitates various chemical reactions through the concentration of light and hot-carrier dynamics (Fig. 3). The first mechanism involves the extreme concentration of light, which significantly promotes the excitation of electrons within the adsorbed molecule. Such an excitation takes place within the adsorbate only but is significantly enhanced by plasmonics, which is referred to as the enhanced intramolecular excitation mechanism.\textsuperscript{111,113,160} The second mechanism involves the transfer of plasmonic hot carriers. Due to the hybridization of the adsorbate–nanoparticle system, the generated hot electrons (or holes) can transfer to the adsorbed molecules. Such transfers result in the placement of an adsorbate–nanoparticle system onto a manifold of excited states since the indirect transfer mechanism involves the enhanced intramolecular excitation mechanism.\textsuperscript{114,164,165} Third, due to the hybridization, direct excitation of electrons from metal states near the Fermi level to the unoccupied molecular orbitals (LUMO) of the adsorbed molecule can occur when the plasmon frequency is resonant with the excitation energy between the metallic occupied orbitals and hydrated metal–molecular orbitals. Such a reaction pathway is referred to as the direct charge transfer (CT) mechanism. Such excitations circumvent the metal–molecular hybridization, direct excitation of electrons from metal states to the adsorbed molecule.\textsuperscript{22,112} Finally, local heating resulting from hot-carrier relaxation can thermally activate a reaction.\textsuperscript{114,164,165}

These mechanisms share some common features that render their clear distinction very challenging. Indeed, different mechanisms have been proposed even for the same chemical reaction, leaving a very confusing situation. For example, The chemical decomposition of methylene blue molecules on the surface of plasmonic NPs was reported by different research groups,\textsuperscript{115,160,161,166,167} direct hot electron transfer mechanism was proposed in ref. 115 and 161 and indirect hot electron transfer is proposed in ref. 167. The same mechanism also applies to the plasmon-mediated reduction of PNTP, where both the nonthermal and local heating effects were proposed.\textsuperscript{168,169} Hence, in order to obtain a comprehensive understanding of plasmon-mediated chemistry, it’s essential for the theoretical methods to fulfill the following requirements:

1. Efficient computation of a dense manifold of excited states because plasmonic excitations in metallic nanostructures are usually not low-lying states. Even for a small nanoparticle (2 nm diameter, for example), thousands of excited states are to be computed to reach the plasmon excitations.\textsuperscript{108}

2. The ability to capture the nonadiabatic transition between excited states since the indirect transfer mechanism involves the transition between the hot electron state (excitations localized within the plasmonic system only) and the charge transfer state.\textsuperscript{110}

3. Inclusion of electron-vibrational couplings that lead to the hot electron relaxation and local heating.

4. Inclusion of electron–electron scattering that leads to the redistribution of hot electrons.\textsuperscript{104,170}

In particular, recent debates on the thermal impact underpin the necessity of atomistic and dynamical insights via nonadiabatic simulations of the HE generation, transfer, and relaxation processes on equal footing.

### 3.5 Nonadiabatic simulation of plasmon-mediated chemical reactivities

Ehrenfest dynamics and Surface Hopping. With the Born–Oppenheimer approximation, the electronic wave function \( \Theta(\mathbf{r}, \mathbf{R}) \) is expanded on the basis of adiabatic BO states, which depend on the electronic coordinates \( \mathbf{r} \) and the nuclear coordinates \( \mathbf{R}(t) \) according to

\[
\Theta(\mathbf{r}, \mathbf{R}) = \sum_{n=1}^{N_a} c_n(t) | \phi_n(\mathbf{r}, \mathbf{R}(t)) \rangle.
\]  

(31)

Here \( N_a \) is the total number of adiabatic electronic states, \( | \phi_n(\mathbf{r}, \mathbf{R}(t)) \rangle \) is the adiabatic electronic wavefunction of state \( n \) and \( c_n(t) \) are the time-dependent complex expansion coefficients. \( \mathbf{R}(t) = \{ \mathbf{R}_A(t) \} \}_{i=1}^{N_a} \) (\( N_a \) is the total number of atoms in the system) are the nuclear trajectories which are obtained by solving the classical Newton’s equations of motion (EOMs) with the mixed quantum-classical framework.\textsuperscript{171}

\[
M_A \frac{d^2 \mathbf{R}_A}{dt^2} = -\nabla_{\mathbf{R}_A} E(\mathbf{R}),
\]

(32)

where \( M_A \) is the mass of \( A \)th atom. \( E(\mathbf{R}) \) is the potential energy surface (PES), which can be an averaged one \( E(\mathbf{R}) = \sum_n c_n(t) E_n(\mathbf{R}) \) within the mean-field Ehrenfest dynamics or a single PES of a certain state \( E_n(\mathbf{R}) \) within the Surface hopping (SH) framework.\textsuperscript{171–173}

By substituting eqn (31) into the time-dependent Schrödinger equation and keeping only the first-order nonadiabatic coupling terms, a set of EOMs for the coefficients \( c_n(t) \) along a given classical trajectory can be obtained\textsuperscript{171,174}

\[
i \hbar \frac{\partial c_n(t)}{\partial t} = c_n(t) \dot{E}_n(\mathbf{R}) - i \hbar \sum_m c_m(t) \nabla_{\mathbf{R}} \cdot \mathbf{d}_{nm}.
\]

(33)

Here the orthogonal condition of adiabatic states \( \langle \Psi_n | \Psi_m \rangle = \delta_{nm} \) is used. \( \mathbf{d}_{nm} = \langle \Psi_n | \nabla_{\mathbf{R}} \Psi_m \rangle \) is the nonadiabatic derivative.

---

**Fig. 3** Proposed mechanisms of plasmon-mediated reactivities. (1) Enhanced intramolecular excitation, (2) indirect charge transfer, (3) direct charge transfer, and (4) local heating.
coupling term (or nonadiabatic coupling vector, NACR). A key variable in eqn (33) is the time-derivative nonadiabatic coupling scalar (NACT) between two adiabatic states

$$\textbf{R} \cdot d_{nm} = \left< \psi_n \left| \frac{\partial}{\partial t} \right| \psi_m \right>, \quad (34)$$

which is responsible for the nonadiabatic transitions between different adiabatic states and can be easily calculated with many \textit{ab initio} methods for excited states.

With the SH algorithm, the state that governs the dynamics of nuclei is determined by a stochastic process. The time-dependent coefficients \( c_n(t) \) obtained in eqn (33) are used to calculate the hopping probabilities between different electronic excited states within the framework of the FSSH algorithm. The hopping probabilities between excited states \( n \) and \( m \) are given by\(^{175}\)

$$g_{n \rightarrow m}(\textbf{R}, t) = \frac{1}{t} \frac{\int_{t}^{t+\Delta t} \text{d}t \, b_{nm}(\textbf{R}, t)}{d_{nm}(t)}, \quad (35)$$

where \( \Delta t = \frac{\Delta \xi}{\dot{\xi} t} \) with \( \Delta \xi \) and \( \dot{\xi} t \) correspond to the step sizes for evolving motions of nuclei and electrons in eqn (32) and (33), respectively. The chosen value of \( \Delta t \) must be small enough to resolve strongly localized peaks in NACT in order to avoid underestimation of transition probabilities. This is particularly important when crossings between adiabatic states are encountered in the trajectory, in which \( \Delta t \) will be further refined. \( a_{nm}(t) = c_n(t)c_m^*(t) \) defines the time-dependent density matrix elements, and \( b_{nm}(\textbf{R}, t) = -2\text{Re}(a_{nm}^\ast \textbf{R} \cdot d_{nm}) \). Note that \( g_{n \rightarrow m} = -g_{m \rightarrow n} \) and \( g_{n \rightarrow n} = 0 \) since \( d_{nm} \) are antisymmetric. Hopping between adiabatic states is determined stochastically by comparing \( g_{n \rightarrow m} \) to a random number \( \xi(\xi \in (0, 1)) \). A hop from state \( n \) to state \( m \) is performed if

$$\sum_{l=1}^{m-1} g_{n \rightarrow l} < \xi \leq \sum_{l=1}^{m} g_{n \rightarrow l}, \quad (36)$$

where states are assumed to be ordered with increasing transition energy. On the other hand, the system remains in state \( n \) when \( \sum_{l=1}^{N_n} g_{n \rightarrow l} < \xi \). If \( g_{n \rightarrow m} < 0 \), the hop is unphysical, and the probability is set to zero. Finally, if a hop to a higher energy state is predicted, there must be sufficient nuclear kinetic energy along the direction of NACR. Otherwise, the hop is rejected. After a successful hop, the total electron-nuclear energy is conserved by rescaling the nuclear velocity in the direction of the NACR according to the procedure described in ref. 173 and 176. In addition, during the dynamics, we monitor the relative phase of the ground to excited state transitions and maintain the same phase (sign) to avoid a sudden sign change in the NACT. This is done by enforcing the sign of the largest component of the Casida eigenvectors to the same along the trajectory.

Despite broad popularity in the community, either Ehrenfest dynamics or the surface hopping (SH) approach have well-known limitations such as generating artificial electronic coherence or giving incorrect long-time population.\(^{173}\) To address these challenges and provide accurate dynamics, a multiconfigurational Ehrenfest (MCE) dynamics approach\(^{177}\) and \textit{ab initio} multiple cloning (AICM)\(^{178-181}\) are developed accordingly.

Multiconfigurational Ehrenfest (MCE). MCE generalizes EHR formalism by representing the wave function as a linear combination of Ehrenfest configurations. Each MCE configuration moves along its own Ehrenfest (mean-field) trajectory. Within the MCE formalism, the molecular wavefunction \( |\psi\rangle \) is expressed in the trajectory-guided Gaussian basis functions (TBF) representation (|\( \psi_n \rangle \)),

$$|\psi(t)\rangle = \sum_n c_n |\psi_n(t)\rangle. \quad (37)$$

And each configuration (or TBF) is described by the product of nuclear and electronic parts,

$$|\psi_n(t)\rangle = |\chi_n(R(t))\rangle \sum_{\alpha} d^{\alpha}_{n} |\phi^{\alpha}_{\alpha}(r, \textbf{R}(t))\rangle. \quad (38)$$

Where \( |\phi^{\gamma}_{\alpha}(\textbf{r}, \textbf{R}(t))\rangle \) is the adiabatic state of configuration \( n \), \( |\chi_n\rangle \) are Gaussian nuclear basis functions

$$|\chi_n\rangle = \left( \frac{2\pi}{\alpha} \right)^{N_a/4} e^{-\frac{1}{2} \alpha \left( \xi(R-R)^T + \frac{\beta}{2} (\xi(R-R) + \xi(R-R)) \right)}. \quad (39)$$

The couplings between TBFs in the MCE approach are described by the EOM of \( c_n(t) \), which can be readily obtained by substituting eqn (37) into the Schrödinger equation:

$$i\hbar \sum_n S_{mn} c_n = \sum_n \left[ H_{nm} - i\hbar \left\langle \psi_m \left| \frac{\partial}{\partial t} \right| \psi_n \right\rangle \right] c_n. \quad (40)$$

where

$$H_{nm} = \sum_{IJ} (d^I_n)^* d^J_n \langle \chi^I_m \phi^I_{\alpha} | T + V | \chi^J_m \phi^J_{\alpha} \rangle. \quad (41)$$

and the overlap \( S_{mn} \) is

$$S_{mn} = \left\langle \psi_m | \psi_n \right\rangle = \langle \chi_m | \sum_{IJ} (d^I_n)^* d^J_n \langle \phi^I_{\alpha} | \phi^J_{\alpha} \rangle. \quad (42)$$

The nuclear part of eqn (41) can be obtained analytically,

$$\langle \chi^I_m \phi^I_{\alpha} | V | \chi^J_m \phi^J_{\alpha} \rangle = \langle \chi^I_m | \frac{\hbar^2}{2} \nabla_R M^{-1} \nabla_R | \chi^J_m \rangle \langle \phi^I_{\alpha} | \phi^J_{\alpha} \rangle. \quad (43)$$

While the electronic part (or the potential energy matrix elements) are approximated by\(^{181}\)

$$\langle \chi^I_m \phi^I_{\alpha} | \phi^J_{\beta} \rangle = \frac{1}{2} \langle \phi^I_{\alpha} \rangle \langle \phi^J_{\beta} \rangle \langle \chi^I_m | \chi^J_m \rangle \times \left\{ (V^m_T + V^\alpha_T) + \frac{i}{\hbar} \left[ (\textbf{P}_m - \textbf{P}_\alpha) \cdot (\nabla_R V^m_T + \nabla_R V^\alpha_T) \right) \right. \right.$$

$$- \frac{1}{2} \left( \nabla_R V^m_T - \nabla_R V^\alpha_T \right) \} \right). \quad (44)$$

As shown above, within the MCE schemes, the electronic states are now different for different configurations. The overlaps between TBFs have to be calculated and taken into account.
Ab initio multiple cloning (AIMC). The AIMC method combines the best features of ab initio Multiple Spawning (AIMS) and Multiconfigurational Ehrenfest (MCE) methods. Similar to the MCE method, the individual trajectory basis functions (TBFs) of AIMC follow Ehrenfest equations of motion. However, the basis set is expanded in a similar manner to AIMS when these TBFs become sufficiently mixed. Consequently, AIMC avoids prolonged evolution on the mean-field potential energy surface (PES).

Within the MCE formalism, the Ehrenfest basis set is guided by an average potential, which is accurate for dynamical processes where the coupling between states persists in time between nearly parallel PES. But the mean-field treatment can be unphysical when the PES of two or more populated electronic states become different in shape, which leads to wave packet branching after leaving the nonadiabatic coupling region. To deal with these cases, the AIMC algorithm is applied to expand the original basis set of TBFs by "cloning" one TBF into two copies in a way that does not alter the original wave function. This is done by creating one of the clones in a pure state and the other clone, which includes contributions from all other electronic states. The corresponding MCE amplitudes are adjusted to conserve the original wavefunction.

It should be noted the MQC method can be complemented with any electronic structure solvers as long as the gradients and NACs are available. Practical applications need to balance accuracy and numerical efficiency.

3.6 Simulation of plasmon-mediated phenomena

3.6.1 Jellium model-based NAMD simulation. Compared to the costly atomistic ab initio calculations for plasmonic nanoparticles, the optical absorption of simple sp-metal nanoparticles described by the Jellium model can be easily calculated and analyzed. In general, good agreement with experimental optical properties can be achieved by tuning the Jellium radius. Using the Jellium model, we investigated the generation and relaxation of plasmonic hot carriers, as well as the atomic-scale mechanism of plasmonic hot-carrier-mediated chemical processes such as H2 dissociation. Our numerical simulations showed that after photoexcitation, hot carriers transfer to the antibonding state of the H2 molecule from the nanoparticle, leading to a repulsive-potential-energy surface and H2 dissociation (Fig. 4(b)–(d)). This process occurs when the molecule is close to a single nanoparticle. However, in a plasmonic dimer, dissociation can be inhibited due to sequential charge transfer that effectively reduces the occupation of the antibonding state, as shown in Fig. 4(e). When the molecule is asymmetrically positioned in the gap, the symmetry is broken, and dissociation is restored by significantly suppressing additional charge transfer. Thus, these models illustrate the potential for structurally adjustable photochemistry through plasmonic hot carriers.

3.6.2 TDDFT calculations. However, the Jellium model oversimplifies the electronic structures of the plasmonic nanostuctures, and it lacks atomistic details. Insights from PES built on the ab initio atomistic model are essential for a more in-depth understanding of plasmon chemistry. Thus, we employed the linear response TDDFT (LR-TDDFT) calculations within the Casida formalism to compute the adiabatic PESs of the H2 molecule adsorbed on an Au6 cluster (H2@Au6), in order to explore key pathways in LSPR-promoted chemical reactions. Despite the model system being too small to support plasmonic mode and thus cannot describe the dephasing of plasmons that produce hot electrons, the key point of using the model system is to capture key aspects of the later stages of plasmon-facilitated photocatalysis, thus providing mechanistic insights.

Our findings based on DFT calculations indicate that in the ground state, the Au6 cluster supports the adsorption of the H2 molecule at the tip site. We further use LR-TDDFT calculations within the Casida formalism to determine the adiabatic excited states and corresponding oscillator strengths. We calculate two-dimensional PESs in the desorption and dissociation reaction coordinates and observe that the adiabatic excited PESs bear similarity to the ground state PES in the Franck–Condon region. Therefore, the H2 adsorbate is comparatively stable, as corroborated by the relatively low dissociation and desorption probabilities from our quantum dynamics simulations. By developing the orbital wave function overlapping (OWO) diabatization scheme, we are able to divide the dense manifold of the excited state into two groups: (1) one group is dominated by electronic excitations confined to the Au6 cluster, which can be likened to HE states in metal nanoclusters; (2) the other group of excited states has the antibonding σ* characteristics of the H2 adsorbate due to the hybridization between H2 and Au6 cluster, which are denoted as CT states. The crossings among the HE and CT states provide pathways leading from the excited HE states to CT states via nonadiabatic transitions.
Quantum dynamics simulations on the diabatic PESs demonstrate that the CT diabatic states are able to drive \( \text{H}_2 \) dissociation efficiently and thus are responsible for the experimentally observed HD formation on Au nanoparticles. Our results nevertheless give a clear physical picture of photoinduced \( \text{H}_2 \) dissociation on Au clusters.

The presence of HE-CT crossings is not unique to plasmonic catalysis. Such features have also been found in other non-plasmonic cata
sysis, such as those on the surface of semiconductors.\(^{186}\) However, plasmonic materials are unique in generating a high concentration of HE due to their larger absorption cross sections. The combination of a high concentration of HE and HE-CT crossings is thus the distinctive feature of plasmonic catalysis.

### 3.6.3 TDDFTB-based NAMD simulations

As shown in Fig. 5, plasmon-mediated chemistry involves a dense manifold of excited states, nonadiabatic transitions between HE and CT states, and their competition with various dissipation channels. Such complicated processes require nonadiabatic simulations that treat the HE generation, relaxation, and HE-CT transitions treated on equal footing. To this end, larger clusters should be used in modeling plasmonic catalysis, and more efficient semiempirical methods, such as those based on time-dependent density functional tight binding (TDDFTB),\(^{187}\) might be needed for the efficient electronic structure calculations. To this aim, we recently developed an efficient NAMD method by combining the TSH algorithm and LR-TDDFTB method.\(^{175,188-190}\) LR-TDDFTB is the tight-binding extension of TDDFT,\(^{188,191}\) which has been successfully applied to investigate the optical properties of plasmonic NPs.\(^{192,193}\) It enables calculations of plasmonic excitation along with several hundred excitation states per time step. The relaxation processes of plasmon induced by electron-phonon interactions are treated by the TSH algorithm. With the NAMD-DFTB method, we demonstrate the plasmon relaxation of an Au\(_{20}\) cluster. Our simulations show that Au\(_{20}\) can support plasmon-like excitation, which includes the superposition of multiple single-particle excitation components.

The numerically efficient LR-TDDFTB method allows us to address a dense manifold of excited states to ensure the inclusion of plasmon excitation. Starting from the photoexcited plasmon states in Au\(_{20}\) cluster, we find that the time constant for relaxation from plasmon excited states to the lowest excited states is about 2.7 ps, mainly resulting from a step-wise decay process caused by low-frequency phonons of the Au\(_{20}\) cluster. Furthermore, our simulations show that the lifetime of the phonon-induced plasmon dephasing is \(~10.4\) fs, and such a swift process can be attributed to the strong nonadiabatic effect in small clusters. Our simulations demonstrate a detailed description of the dynamic processes in nanoclusters, including plasmon excitation, hot carrier generation from the plasmon excitation dephasing, and the subsequent phonon-induced relaxation process.

The NEXMD-DFTB method was employed to investigate the plasmon-induced bond activation of CO adsorbed on Au\(_{20}\) cluster.\(^{103}\) The simulations provide a comprehensive and accurate depiction of the multiple dynamic processes involved in plasmon-mediated chemistry, including plasmon excitation, HEs relaxation, direct/indirect HE transfer, and the activation of CO vibration mode induced by HE transfers. These simulations reveal the critical role of charge transfer (CT) states in plasmon-induced CO activation. The CT states excite both direct and indirect HE transfer, which leads to the activation of CO stretching mode, an essential component of plasmon energy relaxation. Our simulations demonstrate the high efficiency of CO vibrational mode activation, achieving a success rate of approximately 40%. Notably, the HE transfer occurs at a faster rate than the conventional scattering process of the Au\(_{20}\), completing within approximately 100 fs, while the energy relaxation takes place over a timescale of approximately 1 ps. Furthermore, our direct atomistic simulations provide detailed insights into the potential energy evolution during the plasmon-mediated chemical transformation, thereby enabling a comprehensive understanding of the energy relaxation and HE transfers during the reaction, as well as elucidating the reaction pathways.

## 4 Polariton chemistry

Since there are several recent experimental and theoretical reviews on polariton chemistry,\(^{25,194-198}\) only a brief introduction to the recent theoretical development of polariton chemistry will be reviewed in this perspective. Instead, we will focus on our perspective on the theoretical challenges and ongoing and future development of quantum many-body and multiscale methods toward solving the mysteries of polariton chemistry and beyond.

### 4.1 Introduction and background

As introduced in Section 2, in the strong light-matter coupling limit (strong enough to compete with individual DOF's
dissipation rate), the formed quasiparticles (polaritons) can alter the potential energy landscape and consequently manipulate the chemical processes. Generally, the strong coupling regime is characterized by the coherent energy exchange between the photon field and the electronic emitter (i.e. Rabi oscillations). In this regime, the electronic and photon subsystems can no longer be treated separately, making the accurate simulation of polariton dynamics very challenging. In this section, we explore various theoretical approaches in modeling the light-matter interactions present in optical cavities, such as Fabry-Pérot, where the coherent exchange in energy between all DOFs directly affects the resulting chemistry. Here, the electronic and photonic DOFs must be treated quantum mechanically on equal footing.

Recent experiments have shown a propensity to change chemistry via the coupling of quantized radiation and various molecular DOFs, namely electronic and vibrational strong coupling. The Fabry–Pérot-like cavities offer an extremely tunable cavity frequency (via the effective length between the cavity mirrors) while exhibiting widely varying coupling strengths highly dependent on the experimental setup. There are many open questions regarding these experiments, such as the collective effects (i.e., many-molecule or many-mode effects), which are present in many of the recent works due to the complexity of performing single-molecule experiments.

### 4.2 Theoretical and computational challenges

There are conceptual and technical challenges in performing simulations of polaritonic systems. Like the situation in plasmon chemistry, the challenges in understanding polariton chemistry still lie in the multiple coherent and dissipative processes across various time/length scales.

1. The first theoretical hurdle is that of the Hamiltonian itself. For the Pauli–Fierz Hamiltonian (eqn (12)), the light-matter interaction requires explicit knowledge of the molecular dipole operator \( \hat{D} \) in the working basis. It turns out that a significant simplification can be made if one neglects all the contributions from the dipole matrix except for the ground-to-excited transition dipole matrix elements. There are two primary reasons for making this approximation: (I) the entirety of this quantity is not usually printed by default when executing standard electronic structure software for electronically excited states; however, the usual information given from these calculations is the ground-to-excited transition dipole moment, \( D_{0p} \), which, for example in linear response calculations (e.g., TD-HF or TD-DFT), is a trivial quantity to achieve since the result of such schemes is the ground-to-excited transition density. (II) Historically, the field of quantum optics was narrowly focused on the light–matter interaction between single atoms and the quantized cavity field. In such cases, the diagonal elements of the dipole operator are zero by construction, and the electronic energy differences between excited states were large, leading to neglecting the excited-to-excited dipole coupling. With these approximations, one would find that a reduced Hamiltonian can be achieved, commonly referred to as the Jaynes-Cumming (JC) Hamiltonian.

2. Generally, the strong coupling relies on a larger number of molecules collectively coupled to a cavity, introducing significant challenges to computing the hybridized polaritonic states of many (10^6 or even larger) molecules in the cavity.

3. While plasmonic cavities are able to achieve strong coupling in a few (or even single) molecule limits, these cavities are highly heterogeneous and dispersive. Consequently, multiple models, dispersion of the cavities, coupling beyond the dipole approximation, and the inherent strong dissipation should be taken into account, introducing different computational complexity. Besides, the molecules inside the nanoplasmonic cavities can also interact with the ones outside the cavity. For example, a recent 2DES investigation of J-aggregates in the plasmonic cavity reveals rich photophysical dynamics between the polaritonic states inside the cavity and uncoupled exciton states outside the cavity, which indicates that the molecules outside the cavity cannot be trivially ignored.

4. The damping of the polariton state is usually ignored or introduced by an empirical lifetime of cavity photon (cavity leakage). In principle, cavity leakage depends on the photonic density of states, which can be computed from the first principles. Besides, the light-matter coupling could affect the photonic DOF and, in return, the photon lifetime. Hence, one open question is to what extent the strong coupling changes the photon lifetime, which requires a multiscale method to address the feedback of molecular systems on the cavities and first-principles calculations of cavity leakage.

5. Polariton-mediated phenomena are fundamentally kinetic procedures involving the dynamic interplay between electronic, nuclear, and photonic DOFs at various time and length scales. Indeed, the accurate simulation of the quantum dynamics of a molecule itself is already a challenging task. The inclusion of photonic DOFs in polariton dynamics adds further complexity. To accurately simulate polariton dynamics, efficient and precise calculations of the gradients and derivative coupling of polariton states are essential. Despite the recent advancements in accurately computing gradients of polariton states, more efforts are needed for large-scale ab initio simulations of polariton dynamics of many molecules.

6. Polariton-mediated phenomena are fundamentally kinetic procedures involving the dynamic interplay between electronic, nuclear, and photonic DOFs at various time and length scales. Indeed, the accurate simulation of the quantum dynamics of a molecule itself is already a challenging task. The inclusion of photonic DOFs in polariton dynamics adds further complexity. To accurately simulate polariton dynamics, efficient and precise calculations of polariton states’ gradients and derivative coupling are essential. Despite the recent advancements in accurately computing gradients of polariton states, more efforts are needed for large-scale ab initio simulations of polariton dynamics of multiple molecules.

Hence, simulating polariton chemistry is non-trivial, and state-of-the-art theoretical models and efficient numerical methods are crucial for understanding the polariton dynamics of many molecules.
4.3 Common approximations toward historic quantum optics models

The PF Hamiltonian (as written in eqn (12)) is derived by applying the dipole approximation, which assumes the wavelength of the cavity fields is substantially larger than the matter system so that the spatial dependence of the transverse fields is neglected. In addition, many of the recent works on the simulation of ab initio polaritons have relied on approximate versions of the PF Hamiltonian that stem from historical applications in the quantum optics community. 212–218 In these heavily approximated Hamiltonians, usually a large truncation of the electronic and photonic subspaces is also performed such that only the ground and a single electronic excited state, \(|g\rangle\) and \(|e\rangle\) are included while only including the vacuum and singly excited Fock states, \(|0\rangle\) and \(|1\rangle\) in a single-mode cavity \(\varepsilon = 0\). The total basis for this simple model is then confined to singly excited Fock states, \(|0\rangle_i^e\) and \(|e\rangle_i^e\) such that only the ground and a single electronic excited state, \(|0\rangle_i\) and \(|e\rangle_i\) are excited. Note here that another approximation is that there is no permanent dipole in the ground \(|g\rangle_i\) and excited \(|e\rangle_i\) electronic states, i.e., \(D_{gg} = D_{ee} = 0\). The most commonly used Hamiltonian for modeling \(ab\) \(initio\) polaritons is the Jaynes–Cummings (JC) Hamiltonian \(\hat{H}_{JC}\), which can be written as,

\[
\hat{H}_{JC} = \hat{H}_m + \hat{H}_p + \sqrt{\frac{\Delta}{2}} \cdot \hat{D} (\hat{\sigma}^+ \hat{a} + \hat{\sigma}^- \hat{a}^\dagger),
\]

where \(\hat{\sigma}^+ (\hat{\sigma}^-)\) is the creation (annihilation) operator for the molecule excitation between the ground \(g\) and excited \(e\) states. And the Tavis–Cummings (TC) \(^{219}\) model is the many molecules generalization of the JC model. \(^{212}\) Besides, the feedback of molecular systems on the cavity modes (via the current density) is usually ignored. Here, two approximations have been made to derive JC/TC models: (I) the rotating wave approximation (RWA) – which is to say, neglecting the highly oscillatory \(\hat{\sigma}^+ \hat{a}^\dagger s\) and \(\hat{\sigma}^- \hat{a} s\) terms in the light–matter interaction – and (II) neglecting the DSE \(\hat{H}_{DSE}\). This Hamiltonian is valid at ultra-low coupling strengths where the splitting between the one-photon-dressed ground state \(|g,1\rangle\) and the excited state with zero photons \(|e,0\rangle\) exhibit linear splitting (e.g., Rabi splitting) with an increase in the coupling strength \(\Delta\). The other two basis states \(|g,0\rangle\) and \(|e,1\rangle\) are completely decoupled from the interaction.

Two other common approximations to the PF Hamiltonian are the Rabi model \(\hat{H}_{Rabi}\) after explicitly dropping the dipole self-energy (DSE) term and the RWA \(\hat{H}_{RWA}\). These two approximate Hamiltonians can be written as,

\[
\hat{H}_{Rabi} = \hat{H}_m + \hat{H}_p + \sqrt{\frac{\Delta}{2}} \cdot \hat{D} (\hat{\sigma}^+ \hat{a}^\dagger + \hat{\sigma}^- \hat{a}), \tag{46}
\]

\[
\hat{H}_{RWA} = \hat{H}_m + \hat{H}_p + \sqrt{\frac{\Delta}{2}} \cdot \hat{D} (\hat{\sigma}_g \hat{a}^\dagger \hat{\sigma}_e \hat{a} + \hat{\sigma}_e \hat{\sigma}_g \hat{a}^\dagger \hat{a}) + \hat{H}_{DSE}. \tag{47}
\]

where the Rabi Hamiltonian \(\hat{H}_{Rabi} = \hat{H}_{PF} - \hat{H}_{DSE}\) and the RWA Hamiltonian applies the RWA to the PF Hamiltonian. Here, \(\hat{\sigma}_g = |g\rangle \langle e|\) is the annihilation operator of the two-level electronic system. A more in-depth discussion on these Hamiltonians and the effects of the dipole self-energy term can be found in ref. 194, 212, 220 and 221.

Since the correct PF Hamiltonian in the dipole gauge has been known, the question remains of why the community returns to the approximated Hamiltonians for \(ab\) \(initio\) as well as model calculations. There are many subtleties to using the full PF Hamiltonian. From the electronic structure perspective, the many-level dipole operator needs to be computed as well as its square. The DSE term provides a very complicated description of the system for large coupling strengths since the dipole matrix, \(\hat{D}\), in realistic molecules is far from sparse with its square leading to further complications.\(^ {57,194}\) This allows for strong coupling between arbitrary states that is not trivial to know \(a\) \(priori\) based on chemical or physical intuition. Often, these Hamiltonians are parameterized based on cavity-free electronic structure calculations to obtain the energies and dipoles of the electronic adiabatic states (see more details in Section 4.4.2). In this case, the number of included electronic (as well as photonic) basis states should be treated as a convergence parameter. In this case, the DSE causes mixing between far-separated-in-energy electronic states, which leads to slow convergence in the basis set size for the matter DOFs, \(^{212}\) and in general poor results compared to benchmarks, especially at large light-matter coupling strengths when using a small electronic basis.\(^ {222–224}\) This is precisely why many authors are developing self-consistent formulations to construct the low-energy polaritonic states without the need to calculate all the high-energy electronic states (see Section 4.4.2).\(^ {57,222,224–229}\)

Further when considering the quantum dynamical propagation of polaritons in \(ab\) \(initio\) systems, additional nuclear gradients are required compared to cavity-free simulations where only the gradients of the adiabatic state energies are required to propagation the quantum dynamics, namely the nuclear gradients on the adiabatic dipole matrix and its square (i.e., \(\nabla_a \hat{D}\) and \(\nabla_a \hat{D}^2\)).\(^ {213}\)

4.4 Theoretical modeling of polariton chemistry

4.4.1 Cavity Born–Oppenheimer approximation. Within the Born–Oppenheimer (BO) approximation,\(^ {230}\) the total electronic–photonic–nuclear can be factorized as,

\[
\Phi(r, R, q_a) = \chi(R) \Psi(r, q_a; R), \tag{48}
\]

where \(\chi(R)\) and \(\Psi(r, q_a; R)\) are the nuclear and polaritonic (i.e., electronic and photonic) wavefunctions, respectively. Note here that the polaritonic wavefunction is parameterized by the nuclear positions, exactly like the case without photonic DOFs outside the cavity. Further, one can invoke the usual Born–Huang-like expansion over the Born–Oppenheimer factorization as,

\[
\Phi(r, q_a) = \sum_{\mu} Z_{\mu}(R) \psi_{\mu}(r, q_a; R), \tag{49}
\]

where \(\psi_{\mu}(r, q_a; R)\) are the BO wavefunctions analogous to those as outputted in standard electronic structure packages for the ground and excited adiabatic states. In this basis, which we will call the adiabatic polaritonic basis to draw a direct connection
to the bare electronic case, we will discuss various ways to calculate such polaritonic wavefunctions \( \psi_{\mu}(r,q;\mathcal{R}) \) from \textit{ab initio} calculations in a variety of approaches and levels of approximation.

In the following three sections, we will explore ways to obtain rigorous nuclear-position-parameterized wavefunctions for the entangled adiabatic electron–photon states \( \psi_{\mu}(r,q;\mathcal{R}) \).

First, a brief description of a direct diagonalization approach (Section 4.4.2) with Hamiltonians parameterized with information from standard electronic structure, while the following sections (Sections 4.4.3 and 4.4.4) will focus on the self-consistent approach toward re-developing the standard many-body schemes in electronic structure theory for the QED Hamiltonian (e.g., QED-HF, QED-DFT, etc.).

4.4.2 Direct diagonalization. The polaritonic state can be readily computed \textit{via} exact diagonalization of polariton Hamiltonian (either JC, TC, Rabi, RWA, or rigorous PF) in a certain basis set. For instance, the matrix elements of the single-mode PF Hamiltonian in the widely used adiabatic-Fock basis can be written as,

\[
\langle \hat{H}_{PF} \rangle_{J,m} = \left[ E_i + \omega_{2} \left( n + \frac{1}{2} \right) \right] \delta_{J,m} + \sqrt{\frac{\omega_{2}}{2}} \hat{\mathbf{D}}_{J} \left( \sqrt{m + 1} \delta_{n,m+1} + \sqrt{m} \delta_{n,m-1} \right) + \frac{1}{2} \sum_{K} \left( \hat{\mathbf{D}}_{JK} \right) \delta_{J,m}.\]

where \( N_{d} \) is the number of adiabatic electronic states included in the basis. It is important to note that in this basis, the PF Hamiltonian is extremely sparse since the coupling elements only connect adjacent Fock states \textit{via} the molecular dipole matrix since the matrix elements of the photonic coordinate \( \hat{q} \) are that of the harmonic oscillator \( (i.e., \text{only have nonzero super- and sub-diagonal elements}) \). Further, the DSE contributions, in general, connect all of the electronic states of the system (with the same photon number) and is the most non-trivial aspect of this Hamiltonian and will vary strongly between molecular systems.

Except for the adiabatic-Fock basis, other options for expanding the JC and PF Hamiltonians exist, such as the coherent states\(^{226,231,232}\) or polarized Fock states.\(^{233}\) In both of these bases, the photonic states are chosen such that the molecular dipole parameterizes the photonic state, thereby, in principle, reducing the convergence of the photonic basis. Further, this “direct diagonalization” approach to solving the PF Hamiltonian is dependent on this basis convergence, and the electronic basis is much more rigid since the adiabatic basis is ubiquitously used for its convenience. However, the convergence of this basis has yet to be thoroughly tested for a wide range of systems in solving the PF Hamiltonian, but it is expected to converge slowly due to the contributions from the DSE term.\(^{57,211}\) This evidences the need to move to a more rigorous self-consistent solution for the polaritonic adiabatic states as is done for the electronic adiabatic states themselves.

4.4.3 Self-consistent polaritonic single-particle approaches \textit{scQED} Hartree–Fock and density functional theories. The mean-field approach to the QED Hamiltonian can be first cast in an identical way as the standard Hartree–Fock procedure in a larger Hilbert space, including the photonic states.\(^{224,228,229,233,234}\) A useful basis, referred to as coherent states (CS), can be performed such that the light–matter interaction part of the PF Hamiltonian can be shifted away.

For a given bare electronic ground state wavefunction \( |\text{HF}\rangle \), the photonic Hamiltonian can be obtained by integrating out the electronic DOF,

\[
\hat{H}_{p} \equiv \langle \text{HF} | \hat{H}_{PF} | \text{HF} \rangle = E_{\text{HF}} + \omega_{2} \left( \hat{a}_{\uparrow} \hat{a}_{\downarrow} + \frac{1}{2} \right) + \sqrt{\frac{\omega_{2}}{2}} \left\langle \hat{a}_{\uparrow} \hat{a}_{\downarrow} + \hat{a}_{\downarrow} \hat{a}_{\uparrow} + \frac{1}{2} \left( \hat{a}_{\uparrow} \hat{a}_{\downarrow} \right)^{2} \right\rangle_{\text{HF}},
\]

where \( \langle \cdots \rangle_{\text{HF}} = \langle \text{HF} | \cdots | \text{HF} \rangle \) is the HF ground state expectation value of the electronic subsystem. The photonic Hamiltonian can be trivially diagonalized by introducing the CS transformation \( \hat{U}(\mathbf{z}) \hat{H}_{p} \hat{U}^{\dagger}(\mathbf{z}) \),

\[
\hat{U}(\mathbf{z}) = e^{i \hat{a}_{\uparrow} \hat{a}_{\downarrow}} \hat{z}_{\downarrow},
\]

where \( \mathbf{z} = (\mathbf{z}_{\downarrow}) \) is a vector of complex numbers specific to each cavity mode,

\[
\mathbf{z} \rightarrow \langle \hat{a}_{\uparrow} \hat{a}_{\downarrow} \rangle_{\text{HF}} \frac{1}{\sqrt{2 \omega_{2}}},
\]

Thus, considering a QED-HF ansatz as \( |\text{HF}\rangle \otimes \hat{U}(\mathbf{z})|0\rangle \), the QED-HF energy and corresponding Fock matrix can be derived. Compared to the HF theory for bare electrons, the QED-HF method introduces DSE-mediated one-body and two-body integrals.\(^{195}\)

Similarly to the \textit{scQED-HF} approach, the self-consistent QED density functional theory (\textit{scQED-DFT}) is composed in a similar way, where the main elements of DFT remain, such as the exchange–correlation functional of the density. There are many ways to set up the \textit{scQED-DFT} problem, such as employing a novel exchange–correlation functional to account for light–matter correlation effects\(^{235–238}\) or working with electron-only exchange–correlation functionals using the coherent state basis for the photonic DOFs.\(^{212,225,226}\) In any case, the resulting ground state is uncorrelated by the nature of the DFT formalism.

Excited states \textit{scQED-TD-(HF,DFT)}. The time-dependent analogues to the aforementioned single-particle approaches are powerful tools to probe non-equilibrium densities that give rise to electronic excited states. One of the most popular approaches is one of linear response (LR), resulting in LR-TD-HF and LR-TD-DFT for bare molecular systems in the random phase approximation (RPA). Although, it should be noted that the real-time propagation of the single-particle density matrix – leading to the real-time TD-HF and real-time TD-DFT approaches – is, in principle, a more robust approach but one that is usually more costly than that of linear response. Such
schemes have already been developed for the simulation of molecular polaritons using classical photon DOFs.\textsuperscript{239,240} Here, we will focus our attention on the LR formalism, specifically using a Casida-like approach to writing the random phase approximation (RPA), originally formulated by Flick and coworkers\textsuperscript{222} using the QEDFT (or scQED-DFT in the notation of this work) method in the language of Casida and further used by the groups of Shao\textsuperscript{212} and DePrince.\textsuperscript{225} It should be noted that other formulations of CIS-like excited states can be found in the community, such as the non-Hermitian CIS aimed at simulating cavity loss via a complex photon frequency.\textsuperscript{224}

As per usual, and following the notation of ref. 225, the LR-TD-HF and LR-TD-HF eigenvalue equations using the Casida formalism can be written to include the QED components that satisfy the Pauli-Fierz Hamiltonian (eqn (12)).

An important distinction between various implementations of the QED-TD-DFT approaches in the community is whether the single-particle orbitals used in the formation are “relaxed” in the presence of the cavity or are simply the bare electronic single-particle states. For example, in ref. 212, the orbitals are not relaxed while in the ref. 222 and 225 the orbitals are relaxed. While it is clear that using a relaxed reference state for the basis of the RPA equations would provide a more rigorous result, it is not clear whether identical results can be obtained in the infinite basis limit of both approaches, i.e., including more single-particle states in the CIS-like expansion in excited Slater determinants. Since the RPA equations are iteratively solved, the expansion coefficients of the excited Slater determinants may result in the same excited state observables in the infinite basis limit, while for a finite basis, it may not.

4.4.4 Self-consistent polaritonic coupled cluster approaches (CC,EOM-CC). Despite computational efficiency, DFT or mean-field HF methods usually underestimate the correlations. In particular, the mean-field method cannot describe the electron–photon and photon-mediated electron–electron correlations, and the exchange–correlation function for electron–photon interaction is unknown. Consequently, the QED counterpart of coupled-cluster theory (QED-CC) is proposed.\textsuperscript{36,37,227–229,241,242} Similar to the conventional CC theory, QED-CC employs an exponential wavefunction Ansatz to derive the ground state,

\[
|\Psi_{\text{CC}}\rangle = e^{\hat{T}}|\Phi_0\rangle,
\]

where $|\Phi_0\rangle$ is the reference wave function, which is usually chosen to be the tensor product of Hartree–Fock (HF) determinant and photon vacuum state, i.e., $|\Phi_0\rangle = |\phi_0\rangle \otimes |\gamma_0\rangle$. $\hat{T}$ is defined in excitation configurations. Within QED-CC theory, the generalized excitation operator includes three components, $\hat{T} = \hat{T}_e + \hat{T}_p + \hat{T}_{ep}$, including electronic ($\hat{T}_e$), photonic ($\hat{T}_p$), and coupled electronic–photonic ($\hat{T}_{ep}$) excitations,

\[
\hat{T}_e = \sum_{ia} \epsilon_i^a \hat{a}_i^\dagger + \sum_{i\alpha j\beta} \gamma_{\alpha\beta}^c \hat{a}_i^\dagger \hat{a}_j \hat{a}_j^\dagger \hat{a}_i \hat{a}_i + \ldots \equiv \sum_{\mu} \tilde{a}_\mu \tilde{a}_\mu^\dagger.
\]

\[
\hat{T}_p = \sum_{a} \gamma_a^p \hat{b}_a^\dagger + \frac{1}{Z} \sum_{\alpha \beta} \gamma_{\alpha\beta}^p \hat{b}_\alpha \hat{b}_\beta + \ldots \equiv \sum_{n} \gamma^n \hat{B}_n.
\]

where $N_e$ and $N_p$ are the numbers of electrons and photons, respectively. $\tilde{a}_\mu = \prod_{\tilde{a}} \tilde{a}_\tilde{a}$ and $\hat{B}_n = \prod_{\alpha} \hat{b}_\alpha$ is the $n$-body photonic excitation. The parameters $(t_{\mu}, n_{\mu}, Z_{\mu, n})$ are the cluster amplitudes. By projecting CC wave-function into a set of orthogonal excited configurations (\{|$\mu$\}),

\[
|\mu\rangle = \mu|\Phi_0\rangle,
\]

Though the scQED-CC theory leads to the FCI solution if no truncation on the excitation operator $\hat{T}$ is applied, the excitation operator is usually truncated at the doubles (CCSD) level in order to trade-off between accuracy and computational efficiency. Some tests have been performed on the level of truncation in the photonic excitations (up to 10) in the CC operator for model systems.\textsuperscript{241} However, the scaling of scQED-CCSD is $C(N_e^6 N_p^4)$ in general, where $M_p$ is the number of photon modes. Such a scaling makes it a bit expensive to rigorously test in extended molecular systems that include many electrons and/or when many photon modes and Fock states are used in the calculations. To further reduce the computational cost, the coherence state\textsuperscript{243} may be used to reduce the photon basis in the scQED-CC calculations.

It should be noted that CC methods exhibit size extensivity even in their truncated forms, meaning that the sum of the energies of non-interacting subsystems is equal to the total energy. This is in contrast to CI approaches, where errors in extensivity can become increasingly large with an increasing number of subsystems. In addition, CC theory has the advantage of size extensivity in that the excitation energies of each subsystem do not vary with the size of the total system when the subsystems are non-interacting, making it an ideal candidate for computing the polaritonic eigenstates of many (non-interacting) molecules.

In addition, the excited polariton states can be computed with the corresponding EOM formalism, which parameterizes a neutral or charged excitation by applying an excitation operator to the CC ground state\textsuperscript{244}

\[
|\tilde{K}\rangle = \tilde{K}|\Phi_0\rangle = \tilde{K}\hat{e}^\dagger|\Phi_0\rangle = E_n\hat{e}^{\dagger}.
\]
Here, $E_n$ is the energy of the $n$th excited state, and $\mu$ indexes an element of the excitation operator $\hat{R}$. The excitation operator, $\hat{R}$, can be chosen to access charged or neutral excitations.

We recently implemented such QED-CCSD/EOM-CCSD method to compute the polariton states. With the diagrammatic technique, an auto code generator and optimizer are developed to generate the QED-CCSD/EOM-CCSD equations to arbitrary photon order. Fig. 6 is an example of the Li–H bond dissociation curves in a single-mode optical cavity. Fig. 6a shows the ground state Born–Oppenheimer potential energy surface for the bare electronic system (black circles) and for the polaritonic system (red curve) at zero light-matter coupling strength ($\lambda = 0$). The excited states inside (red curves) and outside (black circles) of the cavity are shown at zero coupling strength ($\lambda = 0$), where two multiplicities (singlet $|S_1\rangle \equiv |S_{1,0}\rangle$ and triplet $|T_1\rangle \equiv |T_{1,0}\rangle$) of electronic states are shown as well as a single cavity state ($|\omega\rangle \equiv |S_{0,1}\rangle$). Here, we have used the notation $|S_j\rangle \otimes |n\rangle = |S_j,n\rangle$ where the left ket in the product signifies the electronic DOFs while the right signifies the photonic Fock state label (i.e., the number of photons in the photon-dressed electronic state). At finite light-matter coupling of $\lambda = 0.088$ a.u., the cavity state with one photon $|\omega\rangle$ couples strongly with the singlet excited electronic state with zero photons $|S_1\rangle$ while the excited electronic triplet state $|T_1\rangle$ is negligibly affected. The so-called Rabi splitting appears at the degenerate point between the singlet electronic state and the cavity state, $R_{Li-H} \approx 3.75$, forming the upper ($|P_0\rangle$) and lower ($|P_0\rangle$) polaritonic states with mixed electronic and photonic character.

The computational efficiency of many-body electronic (or polaritonic) structure codes is nearly as important as the method itself. We have implemented various backend options for our code (presented in Fig. 6). The efficiency of Numpy (CPU-accessible) and Torch einsum (GPU-accessible) functions are shown in Fig. 7 on a vertical log scale as a function of the number of orbitals (i.e., electrons) included in the calculation. The GPU hardware allows the QED-CC code to consistently operate with an order of magnitude less wall time than the CPU version. These results indicate that the conversion toward GPU hardware is required for exploring the frontier science of molecular polaritons.

5 Perspective and summary

5.1 Multiscale method for light–matter interaction

Though there have been extensive developments in QED methods, most of the current implementations are based on the dipole approximation. However, the dipole approximation can fail in many cases. The nanoplasmonic cavities allow for light localization into deeply subwavelength dimensions, leading to effective mode volumes as small as a few nm$^3$ or even Å$^3$ (picocavities). Consequently, the size of molecules becomes comparable to the cavity volumes, and the widely used dipole approximation breaks down. Moreover, the cavity confines the photon field in a certain direction when coupling molecules with many modes inside a cavity (either nanophotonic or nanoplasmonic ones). For a given probing angle $\theta$ (relative to the normal vector of the cavity), the photon energy has a certain
dispersion function. Under this situation, the dipole approxima-
tion along the in-plane direction of the cavity no longer holds as well.\textsuperscript{214,247} Hence, a general method that can compute the multiple spatial-dependent cavity eigenmodes $\lambda_{nk}(r)$ and light-matter couplings (from Maxwell’s equations) is required.

As argued above, the EM field within the cavity is, in principle, not homogeneous, and spatial dependence matters in light-matter interactions. The spatial distribution of the EM field can be solved \textit{via} standard computational electromagnetic methods, such as Finite-difference time-domain (FDTD)\textsuperscript{246} for Maxwell’s equations. Consequently, the light-matter coupling strength can stem from a delicate balance between the spatial dependence of the electronic wavefunctions and the photonic fields. Therefore, only a quantum model that fully incorporates the inhomogeneities of the exciton transition charge density can quantitatively describe this interplay. Using a fully first-principles methodology to describe the quantum chemistry of molecules placed inside the cavity, we can reveal the limitations of the point-dipole approach to address the exciton dynamics in weak and strong coupling regimes. Besides, current methods usually don’t consider the feedback of molecular systems on the cavities. In many situations, modeling of molecular systems only is insufficient as the molecular response can significantly affect the EM distributions, especially in nanoplasmonic cavities. Even though the \textit{ab initio} QED methods we previously developed take into account the interplay between the electronic and photonic DOFs, the molecular response on the EM environment is not considered.

In principle, we could solve the Schrödinger and Maxwell equations simultaneously in order to obtain access to the radiated fields and, with it to the self-consistent evolution of light and matter, \textit{i.e.}, the cavity photon modes can be affected by the modular dipoles due to the Ampere’s Law (Helmholtz’s equation):

$$\nabla \times \frac{1}{\mu_i(r_0)} \nabla \times -\omega^2 \mu_{0\omega} \varepsilon(r_0) J(r) = -e J(r), \quad (62)$$

$$\left( \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) A = -\mu_0 J. \quad (63)$$

Where $J$ is the paramagnetic current density due to the molecular dipoles. According to Maxwell’s equations, each current induces an electromagnetic field for which its precise spatial and polarization structure depends on the electromagnetic environment—oscillating charges emit light. Hence, a fully self-consistent QED method should consider the feedback of molecular dipoles on the cavity properties. The flowchart is demonstrated in Fig. 8.

Alternatively, the generated field can be expressed with the help of the dyadic Green’s tensor,

$$E(r, \omega) = i\hbar \omega \int \! d^3r' G(r, r', \omega) [-e J(r', \omega)], \quad (64)$$

where Green’s function is the formal solution of Helmholtz’s equation

$$\nabla \times \frac{1}{\mu_i(r_0)} \nabla \times -\omega^2 \mu_{0\omega} \varepsilon(r_0) G(r, r', \omega) = \delta(r, r'), \quad (65)$$

with linear media $\varepsilon(r)$. The Green’s function in eqn (65) can be solved from the Frequency domain FDTD methods.

Utilizing Dyadic Green’s functions offers a powerful approach to simplify the description of certain components of a matter system while focusing on the computation of electronic structure only. Such a concept has been proposed in the past in a semiclassical treatment of light–matter interaction.\textsuperscript{247,248} Introducing Green’s function is particularly advantageous in multicomponent systems that span different length scales, such as a microscopic molecule (which is described by the current density $J$) and macroscopic solvent (represented by a parameterized dielectric function $\varepsilon(r)$). The EM environment embedded \textit{via} Dyadic Green’s function can be obtained numerically using many standard Maxwell solvers with certain boundary conditions, such as FDTD\textsuperscript{246} and method of moments,\textsuperscript{249} providing the photon mode structures that are coupled with QED electronic structure solvers. By computing $G$ beforehand, the self-consistency is embedded solely \textit{via} the current density, eliminating the need to treat QED electronic structures and Maxwell’s equation simultaneously. The bypassing of the self-consistency between Maxwell–Schrödinger equations is a significant strength of Dyadic Green’s function approach. This makes the approach more accessible and easier to implement in various electronic structure solvers. But, it should be noted that introducing linear media restricts the evaluation in the linear response regime.

Moreover, the multiscale methods described above can have a significant influence other than polariton chemistry. It could affect the study of light-matter interactions in chemistry, physics, materials science, and energy science at large, such as plasmon chemistry,\textsuperscript{21} quantum plasmonics,\textsuperscript{86} quantum information transduction,\textsuperscript{250} and photonic-microelectronics integration and polaritonic devices.\textsuperscript{251} The integration with open quantum system method, particularly time-dependent quantum transport theories\textsuperscript{252–254} will also open the door to simulating light-driven quantum transport phenomena\textsuperscript{255,256} in either weak or strong coupling regimes.
5.2 Summary

Controlling chemistry or molecular properties has been a longstanding Holy Grail for many decades. Only recently, new possibilities have emerged in the context of light-matter interactions via either plasmon- or polariton-mediated chemistry, which holds the promise of providing fundamentally new strategies to control chemical reactions that are completely distinct from traditional ones (such as electrochemistry, thermochemistry, and photochemistry). Consequently, the application of light-matter interactions in manipulating chemistry has attracted increasing experimental and theoretical attention. However, the inherent multiscale nature of light–matter interaction problems poses significant challenges for both experimental and theoretical investigations of the underlying processes. Besides, multiscale processes at different length and time scales make it difficult to optimize the performance of light-matter interaction-mediated chemistry, and precise modeling of these processes is essential for obtaining a comprehensive understanding of the fundamental mechanisms or current experiments. Hence, the development of multiscale theoretical and computational models that can precisely describe the dynamical processes in light–matter interaction-mediated chemistry across different time and length scales, in conjunction with massively parallel algorithms for large-scale simulations, is essential for obtaining comprehensive insights into the fundamental mechanisms of current experiments. These simulations allow for the modeling of complex, multi-scale processes that would be otherwise difficult or impossible to observe experimentally alone. Such development will ultimately facilitate the discovery of new reaction pathways enhanced by light-matter interactions.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors acknowledge support from the Laboratory Directed Research and Development (LDRD) program of Los Alamos National Laboratory (Grant No. 20220527ECR) and the support from the US DOE, Office of Science, Basic Energy Sciences, Chemical Sciences, Geosciences, and Biosciences Division under Triad National Security, LLC (“Triad”) contract Grant 89233218CNA000001 (FWP: LANLEC7). LANL is operated by Triad National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy (contract no. 89233218CNA000001).

Notes and references