Unraveling energy conversion modeling in the intrinsic persistent upconverted luminescence of solids: a study of native point defects in antiferromagnetic Er₂O₃

Bolong Huang

We investigated the mechanism of the intrinsic persistent luminescence of Er₂O₃ in the A-type lattice based on first-principles calculations. We found that the native point defects were engaged in mutual subtle interactions in the form of chemical reactions between different charge states. The release of energy related to lattice distortion facilitates the conversion of energy for electrons to be transported between the valence band and the trap levels or even between the deep trap levels so as to generate persistent luminescence. The defect transitions that take place along the zero-phonon line release energy to enable optical transitions, with the exact amount of negative effective correlation energy determined by the lattice distortions. Our calculations on the thermodynamic transition levels confirm that both the visible and NIR experimentally observed intrinsic persistent luminescence (phosphor or afterglow) are related to the thermodynamic transition levels of oxygen-related defects, and the thermodynamic transition levels within different charge states for these defects are independent of the chemical potentials of the given species. Lattice distortion defects such as anion Frenkel (a-Fr) pair defects play an important role in transporting O-related defects between different lattice sites. To obtain red persistent luminescence that matches the biological therapeutic window, it is suggested to increase the electron transition levels between high-coordinated O vacancies and related metastable a-Fr defects; a close-packed core–shell structure is required to quench low-coordinated O-related defects so as to reduce the green band luminescence. We further established a conversed chain reaction (CCR) model to interpret the energy conversion process of persistent luminescence in terms of the inter-reactions of native point defects between different charge states. It is advantageous to use the study of defect levels combined with formation energies to suggest limits to doping energy and explain photostimulated luminescence in terms of native point defects.

Introduction

Persistent luminescence refers to an optical phenomenon in which a material is excited by high-energy radiation such as UV light and the resulting visible luminescence emission remains visible for an appreciable time, from minutes to hours after the stoppage of the excitation. It exhibits greener and more efficient features than conventional luminescence processes. It shows the most promise for applications in bioimaging because no external excitation is needed on tissue. Current demands from biological and medical scientists are for red emissions (therapeutic window in medicine: 600–700 nm) and even emissions in the NIR region (highly tissue-penetrating light: 700–1000 nm), which are optimal for the diagnosis and treatment of living tumor cells via in vivo imaging technology.

Recent progress implies that lanthanide oxide materials are novel upconversion luminescence materials that hold great potential for persistent upconverted luminescence. It is uncommon that lanthanide sesquioxides themselves display persistent luminescence, especially based on the upconversion character of electron transport. Tanner et al. concluded that this property is not due to unique electronic characteristics, as in the case of the 4f-5d or 4f-4f transitions in Er₂O₃, because the experimentally reported persistent luminescence emission bands appear to be the same for different Ln₂O₃ (Ln = Pr, Nd, Sm and Er). Accordingly, this persistent luminescence is attributed to subtle interactions between different states of native point defects, based on the experience of our previous investigations. Unfortunately, the relationship between electron transport assisted by the levels of native defects and
the properties of persistent upconverted luminescence is still far from being fully understood, especially at the electronic level. This knowledge gap not only impedes the process of designing new generations of persistent luminescence devices but also prevents the further development of existing technology. Further investigations to fill this gap are of great significance for the relevant science and related technological applications.

In persistent luminescence, the driving force for extensive theoretical investigations is related to the current consensus as summarized by experiments, in which two dominant factors determine the performance of persistent luminescent materials: activators and traps. The emission wavelength is mainly determined by the intraconfigurational levels provided by activators. The intensity and duration of persistent luminescence are governed by trap centers in terms of native point defects in the lattice or co-dopants. The exact balance of activators and traps within the electronic structure decides the selection of their optimal combinations.

However, some recent experiments have refreshed our understanding of the role of trap levels in persistent luminescence and given rise to our new interest in the native point defects of luminescent solids. In an experiment that controlled trap levels carried out by Maldiney et al., measurements of thermal glow (or thermally stimulated luminescence, TSL) were benchmarked with respect to the trap levels provided by various rare earth (RE) ion dopants. The deepest trap levels, which were induced by Dy³⁺, did not give rise to the lowest intensity. In contrast, Pr³⁺, for which the levels were the second closest to the conduction band edge, provided the longest duration and maintained the highest intensity, where the trap levels were even higher than the levels of native O vacancies. Therefore, TSL measurements under conditions of varying temperature cannot give a perfect quantitative description of the levels of native point defects and related variations in behavior for persistent phosphors. This arises because native point defects in specific charge states may be modified by the external thermal conditions during TSL. This consideration provides the idea that types of native defect such as Schottky (in the rock salt simple cubic lattice) or Frenkel (in the fluorite lattice) pair defects play a significant role that is independent of the chemical potentials in a thermal or optical disequilibrium state with charge neutrality maintained.

The interaction of the levels of different point defects is a significant issue. As known from previous work, we found that the native point defects in the host materials not only produce local lattice distortion or charge density anomalies, but also induce unique localized electron levels in the fundamental optical band gap that are either occupied or empty. Therefore, the stabilities and concentrations of the native point defects in the host materials are quite important in studies of persistent phosphors. Recent work performed by De Vos et al. has provided supportive information via calculations using local density formalism. They found that antisite defects have greater stability and their distance (with respect to the density) is closely correlated to the overall stability of the material based on Zn–Ga oxide alloys (ZnGa₂O₄), which have also been investigated by Karazhanov et al. using DFT calculations. Their localized gap states exhibit a qualitative trend to support their model.

The actual situations regarding the properties of the materials are complicated and it is also necessary to consider the participation of all possible defects and their contributions to the transition levels. This context can lead us to explain an experiment that shows that the electron transport or photocurrent in a zinc gallate nanowire is very sensitive to the presence of these defects such as oxygen vacancies and interstitials. However, relative errors in calculations may exceed 30%, whereas trends in physical chemistry show the significance of discussing the defect states in the spinel lattice in such defective Zn–Ga oxide alloys in order to understand persistent phosphors from different aspects. A different systematic study by Zeng et al. shows that both anion- and cation-related defects play important roles in accommodating charge carriers for transport and energy conversion to generate persistent luminescence. Therefore, it is necessary to have a preliminary discussion of different native point defects in a simplified model of the A-type lattice of Er₂O₃, to investigate their energy conversion mechanisms, and to provide a useful reference for further experiments and various applications.

Here, we discuss the models of luminescence that relate to F-centers and the energy conversion process, which reflect different optical transitions. As is well known, the optical behavior of typical defects in solids is mostly governed by strong electron–phonon coupling, which leads to a broadened photon–electron transition following an evident Stokes shift between the absorption and emission spectra. Accordingly, another important issue to be solved is the defect reaction model for persistent luminescence, to identify the location of the nearly continuous energy source that generates a persistent optical transition. Actually, if defect-related transitions follow the zero-phonon line (ZPL) between different optical transition states, optical absorption and emission will be accomplished with high efficiency in both optical quantum yield and energy conversion. Our question is what happens if the two optical transition states are replaced by two complementary charged defects, which react via charge transfer and electron transitions in closed cycles along the ZPL. This implies that the lattice distortions during these reactions will support electron transfer between different optical transition levels, ideally with an exact amount of negative effective correlation energy (−U_corr), which is shown in Fig. 1.

Here, we also discuss such native complementary charged point defects, which are not only formed for energy conversion but also act as donor–acceptor pairs (DAP), with a DAP band within the forbidden band gap of the host material, as a paired activator for electron–hole recombination and the release of photons. In addition, these DAP bands actually fix the center of the luminescence band at a specific energy level by means of restricted optical transitions via a closed cycle along the ZPL related to −U_eff. A similar concept has been successfully utilized in other solid functional materials using our theoretical modeling.

With the aim of elucidating the relevant scientific issues via our theoretical methodology, this paper will focus on erbium...
Using the hexagonal (A-type) lattice with the constraint of variations in the lattice on the f-level of Er$^{3+}$ ions, we studied sesquioxides to understand their electronic structures, optical properties, and thermodynamic and optical transition levels. A systematic investigation will start from the inherent principles that govern the interactions among the microscopic structures, levels of native defects, localized states and mechanisms of persistent luminescence.

Calculation setup

As is well known, three stable lattice structures have been observed for lanthanide sesquioxides (Ln$_2$O$_3$), namely, hexagonal (A-type, $P\overline{6}m1$), monoclinic (B-type, C12/m1), and bixbyite (C-type, la3). We chose the antiferromagnetic (AFM) ordered lattice for modeling the structure of Er$_2$O$_3$, as this lattice has a ground state below 4 K (the Neél temperature). This was also done in the effort to save unnecessary computational cost in investigating a trend in the physical chemistry of defect states, as well as in the modeling of energy conversion in such oxides, without making an absolutely conclusive determination.

On the other hand, if we consider nanoparticles in the real world, the coordination of oxygen could be very different. Experimentally reported persistent luminescence is based on powder samples of Er$_2$O$_3$. Therefore, we used the simplest lattice structure with various O coordination numbers and antiferromagnetic behaviors. We thus modeled erbium sesquioxides (Er$_2$O$_3$) using the hexagonal (A-type) lattice with the $P\overline{6}m1$ space group.

Accordingly, in this study of defects, to reduce the influence of variations in the lattice on the f-level of Er$^{3+}$ ions, we constrained our structural model of the host materials to an A-type lattice similar to those of La$_2$O$_3$ or Ce$_2$O$_3$, with similar treatment as in the work done by Gillen et al. and Jiang et al. Note that a similar simplification of the structural model has been utilized in work by one of our previous collaborators in order to investigate the levels of native oxygen-related defects in La$_2$O$_3$ (A-type), Lu$_2$O$_3$ (C-type) and LaLuO$_3$ (distorted C-type) uniformly within a model of an A-type lattice, as well as other similar high-K oxides.

The lattice relaxation of Er$_2$O$_3$ in the ground state was performed at the PBE or PBE+$U$ level by the CASTEP code, because PBE, as the first ‘Jacob’s ladder’ in DFT, has been proven to be reliable enough for the structural relaxation and internal geometrical optimization of solids based on 4f-or 5f-levels, regardless of whether ultrasoft or norm-conserving pseudopotentials are used. Theoretical studies show that pseudopotentials in solids based on f-level have a minor or negligible influence on the lattice parameters in DFT or DFT+$U$. However, Hubbard $U$ parameters induce an extra error in lattices when we need them for improving the accuracy of electronic structures, if we select them carelessly or they are calculated without self-consistent determination. We decided to select five orbital components, namely, 4f, 5s, 5p, 5d, and 6s to represent the valence states of the Er atom. We chose norm-conserving pseudopotentials because in the valence electrons region they can reflect the all-electron behavior with an overlapping integral factor of |S-matrix| = 1, which is evidently different from ultrasoft pseudopotentials. The overlapping integral factor will reduce the contribution of the Hubbard potential corrections when applied to the charge density of the single-electron semi-orbitals. Using the OPIUM code for generation with the recent RRKJ optimization method, we “tailor-made” the norm-conserving pseudopotentials within the framework of the most popular KB (Kleinman–Bylander) projectors and non-linear partial-core corrections, which have won recognition as a result of our theoretical developments. We specifically chose the scalar-relativistic averaging scheme with the aim of dealing with the complicated local hybridization of semi-core 4f-5d orbitals and the relativistic effect of heavy Er$^{3+}$ ions, which also describes the spin-orbit coupling effect well at a numerical level.

With the plane-wave basis set, it is highly recommended to carefully extend the cutoff energy to 750 eV to describe the five different orbital components of the valence electrons of Er$^{3+}$, as well as the strongly localized states induced by the 2p orbitals of O. To guarantee convergence and avoid charge/spin out-of-sync ‘sloshing’, we uniformly chose the ensemble DFT (EDFT) method of Marzari et al. For the k-point sampling mesh, we chose a grid of $7 \times 7 \times 4$ for all related reciprocal space integration for the primitive cell of Er$_2$O$_3$. By the above settings, the convergence tolerance of the calculations of total energy was set to no higher than $5.0 \times 10^{-7}$ eV per atom, and the optimization by Hellmann–Feynman forces in calculations of defects was accomplished to a level lower than 0.01 eV Å$^{-1}$. For calculations of supercell-based defects, a 60-atom $2 \times 2 \times 3$ supercell was used for Er$_2$O$_3$. The Baldereschi special k-point (1/4, 1/4, 0)$^{11}$ with the gamma center off was self-consistently selected. Regarding the geometry relaxation, an algorithm based...
on the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method was used throughout all calculations of bulk and defect supercells.

As we realized from theoretical studies, the electronic states that occur in the fundamental optical gap are less sensitive to the nonlinearity of Hubbard U parameters. Therefore, we chose the rotational-invariant type DFT+U method developed by Anisimov. To reduce the influence of the localized hole states produced by the 2p orbitals of O sites, self-consistently determined Hubbard U potentials were also applied to the O 2p orbitals, which have reached a consensus in many oxide materials. Thus, it is necessary to consider self-energy corrections to both f- and p-orbitals for rare earth oxides. The valence electron configuration for correcting the overlap of valence-core charge density in such materials. Thus, it is necessary to consider self-energy corrections for point defects in CeO$_2$. The valence electron configuration for determining the on-site Coulomb energy of the spurious self-energy of electrons. In previous work, we have established a method for determining the on-site self-energy of electrons and the related relaxation of the wavefunction in orbitals so as to obtain accurate orbital eigenvalues for electronic structures. For details of the process, refer to the previous work. With our self-consistent determination process, the on-site Hubbard U parameters for the 4f orbitals of Er and the 2p orbitals of 6-fold coordinated O and 4-fold coordinated O were found to be 9.72 eV, 5.73 eV and 5.49 eV, respectively. For the pseudopotentials of Er used for calculations, we similarly chose a nonlinear core correction technique for correcting the overlap of valence-core charge density in such heavy-fermion elements; a detailed discussion of such a method has been presented in a previous study of the native point defects in CeO$_2$. The valence electron configuration for the pseudopotentials we chose for the generation was 4f$^{11}$, as the ground-state configuration for Er$^{3+}$. All the DFT+U calculations were performed in a theoretical scheme using the norm-conserving pseudopotentials. This helped us to reflect the all-electron behavior of the valence electrons, especially regarding the subtle effects of the 4f electrons and outer 6s electrons.

For the equation that was used for calculations of the formation energy of defects in different charge states ($q$), an overall supercell was established and remained constantly based on the ground-state relaxed primitive cell, to reduce the side effects of changes in enthalpy due to variations in the cell. The formation energy of a target defect $H_q$ in a specific charge state $q$ can be described via a relation between the positions of the Fermi energy ($E_F$) and the chemical potential $\Delta \mu$ of the species of defect $\alpha$, which is shown as follows:

$$H_q(E_F, \mu) = [E_q - E_H] + q(E_V + \Delta E_F) + \sum x n_x (\mu_x + \Delta \mu_x)$$

(1)

We see that $E_q$ and $E_H$ are the total energies of a relaxed defective lattice in charge state $q$ and an ideal lattice of the host material in the ground state, respectively, $\Delta E_F$ is the change in the Fermi energy with respect to the valence band maximum (VBM, $E_V = 0$), and $n_x$ is the number of atoms of element $\alpha$ chosen as target defect sites; finally, $\mu_x$ is the reference chemical potential, based on the well-established methodology provided by the work of Zunger et al. 13567

**Results and discussion**

**Bulk A-type Er$_2$O$_3$**

The A-type lattice of Er$_2$O$_3$, with the P$\bar{3}$m1 space group has been predicted based on a method used in our previous work. In this symmetry, the high-symmetry points $\Gamma$, $K$, and $M$ are relatively significant in the first Brillouin zone (BZ). The ground-state relaxed structure has lattice parameters of $a = 3.632$ Å and $c = 5.763$ Å, respectively. From Fig. 2, calculations of the band structure show that the band gap of Er$_2$O$_3$ in the A-type lattice is 5.506 eV using GGA+U calculations with the U parameters for O 2p and Er 4f predicted from the previous work. The indirect band gap of 5.506 eV is from $\Gamma$ to $K$, whereas the direct optical transition gap of $\Gamma \rightarrow \Gamma$ is about 5.8 eV, which is close to the average band gap of high-K dielectric oxide materials. Gillen et al. 15 and Jiang et al. 33,34 have performed calculations on lanthanide sesquioxides using different methods, respectively. The HSE functional used by Gillen et al. shows the band gap to be 5.30 eV and Jiang et al. reported a band gap of 5.22 eV by the GW method. These are both 0.2 eV lower than the experimentally reported value of 5.49 eV, whereas ours is close to the experimentally reported 5.5 eV. There is a small increase of 0.2 eV from the experimental result from Prokofiev et al. 58. The width of the valence band (VB) is about 4–5 eV, which consists of the levels of O 2p orbitals. The three occupied 4f levels are located about 2.8 eV, 4.5 eV and 5.8 eV, respectively, below the valence band maximum (VBM, 0 eV in Fig. 2), with an overlap from O 2p levels in the VB. The unoccupied 4f level is around 9 eV above the VBM and remains in the conduction band (CB). The conduction band minimum (CBM) mainly consists of Er 5d orbitals.

Regarding the bulk formation enthalpy of crystalline Er$_5$O$_3$, $2\mu_{\text{tot}} = \Delta H_f$(Er$_5$O$_3$), the experimentally observed formation enthalpy of Er$_5$O$_3$ is reported to lie within the range from –19.67 to –19.73 eV at $T = 298$ K (–1897.8 eV or –1904.2 kJ mol$^{-1}$). Our calculations give –19.17 eV by the GGA+U method ($U_f = 9.72$ eV for 4f of Er, $U_p = 5.49$ eV for 2p of O4, and $U_p = 5.73$ eV for O6)

![Fig. 2 Band structure and TDOS of bulk Er$_2$O$_3$ in A-type lattice.](image-url)
at $T = 0 \text{ K}$ (ground state), which agrees well with the experimental value at $T = 298 \text{ K}$ (with a relative error of only 2.5%), as well as a correct physical trend (with a value about 0.5 eV higher), as the monoclinic phase of Er$_2$O$_3$ at low temperatures (around 300 K) is more stable than the hexagonal phase (A-type lattice) of Er$_2$O$_3$. Our method shows greater advantages than the method of plain GGA calculations.

### Oxygen vacancy (V$_0$)

The V$_0$ is the most common native point defect in oxides, and is usually the initial equilibrium stable state for most oxide materials. Its influence on their electrical and optical properties is rather significant. As many studies of oxides by Lucovsky et al.\cite{Lucovsky} show, the V$_0$ usually induces effects on the O K pre-edge of XAS spectra in optical fundamental absorption experiments, which have been observed in most common oxides such as trigonal HfO$_2$, monoclinic ZrO$_2$, rutile TiO$_2$, bixbyite Lu$_2$O$_3$, SrTiO$_3$, LaTiO$_3$, Gd(Sc, Ti)O$_3$, etc.

We considered the V$_0$ in Er$_2$O$_3$ in charge states of 0, +1, and +2. Considering the coordination of O ions, there are two different O sites in Er$_2$O$_3$: 6-fold coordinated and 4-fold coordinated (O$_6$ and O$_4$, respectively). Their vacancy sites are shown as V$_{O6}$ and V$_{O4}$, respectively. The neutral V$_{O6}$ and V$_{O4}$ both leave two localized electrons at nearby Er sites with antiferromagnetic (AFM) spin behaviors. Fig. 3(a) shows the total density of states (TDOS) of V$_{O6}^0$ (AFM) spin behaviors. The two localized states are 1.89 eV below the CBM with spin-up and -down together, whereas the value is 1.14 eV for V$_{O4}^0$, as shown in Fig. 3(b). In combination with our analysis of the projected DOS (PDOS), the nearby empty Er 5d orbitals accommodate these localized electrons ($\sim 1.25$ e for V$_{O6}^0$) and there is also some extent of hybridization with the first nearest-neighbor 4-fold coordinated O 2p orbitals ($\sim 0.90$ e for V$_{O4}^0$). Thus, it is also shown that the O6 and O4 sites have very different energy levels within the crystal lattice that could localize electrons. These two localized electrons are angular-momentum degenerate, giving distorted s-like wavefunctions. A similar case has been found in our previous work.\cite{Pan}

Both V$_{O6}$ and V$_{O4}$ provide only one excess electron at the O vacancy site but give rise to two localized states within the band gap with one empty state. This arises because the single excess electron occupies the orbital that hybridizes with O 2p and Er 5d exhibits splitting into bonding and anti-bonding states, as shown in Fig. 3(a) and (b), with different energy intervals due to the different on-site Coulomb interactions of O6 and O4 sites.

The localized orbitals of these states are shown in Fig. 3(c) and (d). The filled trap level of V$_{O6}$ lies 2.21 eV below the CBM and 1.16 eV lower than its empty state, whereas for V$_{O4}$ these values are 1.82 eV and 1.24 eV, respectively. Compared with previous work on CeO$_2$ and CaS,\cite{Zeng} we find that for coherent charge-transfer metal oxides (CeO$_2$, TiO$_2$, etc.) the V$_0$ displays paramagnetic behavior, whereas for ionic metal oxides or sulfides such as Er$_2$O$_3$, ZnO, CaS, etc., it exhibits ferromagnetic (FM) behavior owing to the absence of energy degeneracy in the hybridized orbital in the vicinity of anion vacancy sites. From another perspective, the empty state is actually a localized hole trap level in Er$_2$O$_3$ with V$_0$. Electron energy loss spectroscopy (EELS) experiments give evidence that the empty d orbital of the cation tends to hybridize with the s or p orbital of an anion to form an ionic hole trap.\cite{Hao}

The V$_{O6}^0$ and V$_{O4}^0$ revert to the AFM state with two spin-aligned empty states in the gap. This shows that the empty hybridized orbital (2p–5d) has levels with different spins. These are 1.38 eV and 0.70 eV below the CBM for V$_{O6}^{2+}$ and V$_{O4}^{2+}$, respectively. From Fig. 3(a) and (b), we see that the VB edges are almost unchanged within all charge states for both V$_{O6}^0$ and V$_{O4}^0$, whereas the CB edges exhibit evident variations near the CBM, which indicates that the localized electrons left by the O vacancy interact strongly with localized Er 5d orbitals of the host lattice, while having less effect on delocalized 2p orbitals of the host lattice. This implies that the localized electrons left by the O vacancy are predominantly from neighboring Er sites. Fig. 3(c) and (d) show that V$_{O6}^0$ and V$_{O4}^0$ have a minor lattice distortion effect, whereas V$_{O6}^0$ and V$_{O4}^0$ push the six Er sites outward while attracting the nearest O sites inward, and this occurs to a greater extent for V$_{O6}^0$ and V$_{O4}^0$. The orbitals with localized electrons in these charge states of the O vacancy have s-like wavefunctions, whereas the localized hole states display some extent of hybridized Er 5d orbitals.

Considering the formation energies, Fig. 3(e) and (f) show that both V$_{O6}$ and V$_{O4}$ in Er$_2$O$_3$ are defects with positive effective correlation energies (+U$_{eff}$) of 0.99 eV and 0.53 eV, respectively. The role of this effect in persistent luminescence has been discussed in the work of Zeng et al.\cite{Zeng} and our previous work.\cite{Pan} This shows that the process of $2V_0^0 \rightarrow V_0^0 + V_0^{2+}$ is an endothermic chemical reaction. We see from Fig. 3(c) and (d) that each charge state of V$_0$ is followed to a different extent by local structural distortions, but the costs of these thermally ionized V$_0$ defect states cannot be compensated by the lattice distortions. Therefore, they exhibit a positive $U_{eff}$ and V$_0$ is also proved to be a widely existing defect as found by electron spin resonance (ESR) measurements, which has the potential to be an F$^-$ center. V$_{O6}^0$ and V$_{O4}^0$ have very similar formation energies of 7.40 eV and 7.48 eV below the O-rich chemical potential limit, whereas these are as low as 0.05 eV and 0.13 eV, respectively, in Er-rich conditions; this shows a consistent physical trend of rather stable O vacancy defects in high-K oxides such as HfO$_2$ as found by Ka et al.\cite{Ka} The relatively low formation energy indicates that the O vacancy is energetically favorable in erbium sesquioxides, as in deficient erbium oxides Er$_2$O$_3$. The V$_{O6}$ defect has energies of 3.31 eV and $-4.04$ eV below the O-rich and Er-rich limits, respectively. The thermodynamic transition levels of the (0/+1) and (+/2+) states are 4.09 eV and 3.10 eV above the VBM respectively. V$_{O4}$ is shown to be more stable than V$_{O6}$, as their formation energies are 3.10 eV and $-4.25$ eV with transition levels of 4.38 eV and 3.85 eV for the (0/+1) and (+/2+) states, respectively.

As carefully learned from the results of Pan et al.,\cite{Pan} we see that a discussion of the oxygen vacancy is meaningful, as an experiment on Er$_2$O$_3$ by Pan et al.\cite{Pan} shows that the oxygen vacancy controls the electrical conduction, which is related to the transport of charge carriers. More importantly, an experiment by Hao et al.\cite{Hao} shows that the effect of the absorption of water
molecules (H$_2$O) on nanoparticles of Er$_2$O$_3$ will induce the unexpected quenching or lowering of optical excitation transitions, as shown from a comparison of emission spectra between vacuum and ambient conditions (Fig. 2 in ref. 14). Such absorption of moisture is in fact due to the existence of native oxygen vacancy defects in Er$_2$O$_3$, as confirmed by Pan et al.$^{55}$ This effect has also been found by Tang et al. in another oxide system (Bi$_{1-x}$Eu$_x$Ti$_3$O$_{12}$).$^{56}$ On the other hand, Pan et al. have reported further on the effects of oxygen vacancies in other lanthanide sesquioxides such as Lu$_2$O$_3$.$^{67}$ They show that both oxygen vacancies and interstitials (non-lattice defects) are good charge carrier centers.
so as to enable filament conduction in Lu₂O₃.⁶⁷ We similarly deduce that the oxygen-related defects play a similar role in Er₂O₃. Therefore, it is very useful to provide the results of such preliminary theoretical calculations as a reference for their experiments and applications.

O interstitial (Oi)

The Oi in Er₂O₃ produces complicated localized electron and hole states that interact with surrounding lattice sites with different local bonding and charge states. As shown in previous work, the Oi in metal oxides such as CeO₂¹⁷ may form peroxides with O–O homopolar bonds, which gives rise to π-electrons localized along the O–O bonds near Oi sites. The neutral Oi in Er₂O₃ forms an O–O bond and induces four localized π-electron levels below the VBM (from −0.7 to −8 eV), as shown in Fig. 4(a). Other localized occupied states are about 0.2 eV above the VBM. The anti-bonding π-electron states remain at the CBM as π* orbitals of O–O bonds formed by Oi.

The Oi in Er₂O₃ produces clear localized hole states in the band gap 2.51 eV above the VBM, whereas other occupied electron states that are observable remain near the VBM and are 0.2 eV higher. However, these two edge states display a minor ferromagnetic effect as spin-misaligned. The states below the VBM that have shifted into the VB consist of O 2p levels.

The O₂⁻ in Er₂O₃ produces six localized electron states near the VBM with intervals of about 0.15 eV that are spin-aligned. This arises because the Oi accepts two excess electrons and thus gives rise to a 2p⁶ valence electron configuration, in which bonding states are localized near the VBM. We see from Fig. 4(a) that these localized states exhibit antiferromagnetic behavior.

Fig. 4(b) shows the localized electron and hole orbitals near the directions of the O–O bonds formed by the relaxed Oi sites. We see that all the orbitals are localized, which shows the evident trapping behavior exhibited by Oi. The π-electron orbital of Oi is localized around the O–O bond and the nearest-neighbor Oi sites. The Oi produces hybridized pₓ-dᵧ orbitals vertically in the cell along the O–O bond.

Fig. 4(c) shows that the Oi in Er₂O₃ is a defect center with a negative Ueff in the transition state of (2/0), with a value of Ueff of −0.45 eV. This means that the Oi chemical reaction process, 2Oi⁻ → Oi⁰ + O₂⁻, is an exothermic process, and the thermodynamic transition level of (2/0) is 2.16 eV above the VBM. This means that electron spin resonance (ESR) measurements detect a very low density of Oi⁻ defects, because only Oi⁰ and O₂⁻ can stably exist in the lattice and both eigenstates can be stabilized in the gap. The energy cost of the formation of these defects can be compensated by the local lattice distortions.

We see from Fig. 4(c) that below the O-rich potential limit the Oi has a formation energy of 2.16 eV in the Er₂O₃ lattice, whereas the value is 6.48 eV for O₂⁻ below the O-rich limit. The Oi⁻ has an energy cost of 4.55 eV below the O-rich limit, whereas according to the effect of −Ueff this state cannot widely exist in the lattice as it has a low possibility, owing to being constrained by the effect of −Ueff. Thus, Oi is unlikely to form deep hole trap centers in Er₂O₃, at least in the A-type lattice environment.
Fig. 4  (a) TDOS of O$_i$ in neutral (O$^0_i$), singly negative (O$^{-1}_i$), and doubly negative (O$^{-2}_i$) charge states. The dashed line denotes the highest occupied level for electrons. (b) Localized electron and hole orbitals at the relaxed O$_i$ site (Er = green, O = red). (c) Formation energies of O$_i$ below the O- and Er-rich chemical potential limits.
This is consistent with our results deduced from the negative $U_{\text{eff}}$ for the $(-2/0)$ state.

**Er vacancy ($V_{\text{Er}}$)**

The neutral $V_{\text{Er}}$ in Er$_2$O$_3$ ($V^0_{\text{Er}}$) gives rise to three localized holes that occupy the nearby O sites. The single-particle levels (TDOS) shown in Fig. 5(a) confirm that there are three localized deep hole trap levels at about 1.86 eV below the CBM. One of these holes is spin-up, whereas the other two holes remain spin-down. The $V^1_{\text{Er}}$ has two localized hole traps that are 2.88 eV and 2.58 eV below the CBM, respectively, and are spin-misaligned. Fig. 5(b) shows that there are three individual localized empty $p$–$p$ orbitals around the nearest-neighboring O sites as potential electron-accepting centers.

This shows the correct trend of hole passivation by an electron from the CB as the CB edge varies, displaying ferromagnetic behavior in the TDOS (Fig. 5(a)). The $V^2_{\text{Er}}$ exhibits a similar trend, with only one localized hole level at 2.22 eV below the CBM, and the CB displays a flattened behavior with opposite spin states, which indicates the depletion of electrons for hole passivation. Fig. 5(b) shows consistently that two local $p$–$p$ orbitals occupy two of the nearby O sites in the $V^2_{\text{Er}}$ state and only one on-site local $p$–$p$ orbital at an O site in the $V^1_{\text{Er}}$ state.

The $V^3_{\text{Er}}$ has three additional electrons that occupy the VBM with two spins aligned and with only one state observable near the VBM. This means that there are no trap levels produced by this charge state of $V_{\text{Er}}$. No localized hole orbitals are found in Fig. 5(b), as all hole trap levels are occupied by three electrons.

---

**Fig. 5** (a) TDOS of $V_{\text{Er}}$ in neutral ($V^0_{\text{Er}}$), singly negative ($V^1_{\text{Er}}$), doubly negative ($V^2_{\text{Er}}$), and triply negative ($V^3_{\text{Er}}$) charge states. The dashed line denotes the highest occupied level for electrons. (b) Localized hole orbitals at the relaxed $V_{\text{Er}}$ sites and nearby distorted O sites (Er = green, O = red). (c) Formation energy of $V_{\text{Er}}$ below the O- and Er-rich chemical potential limits, displaying a “$-U_{\text{eff}}$” defect with a shallow acceptor trap feature.
From Fig. 5(c), the $V_{\text{Er}}$ is a defect with $-U_{\text{eff}}$ with values of $-2.58$ eV, $-0.48$ eV, $-1.02$ eV, and $-1.56$ eV for the following reactions: $3V_{\text{Er}} \rightarrow 2V_{\text{Er}} + V_{\text{Er}}$, $3V_{\text{Er}} \rightarrow V_{\text{Er}} + 2V_{\text{Er}}$, $V_{\text{Er}} + V_{\text{Er}} \rightarrow V_{\text{Er}} + 2V_{\text{Er}}$, and $2V_{\text{Er}} \rightarrow V_{\text{Er}} + V_{\text{Er}}$, respectively. Thus, we consider that the $V_{\text{Er}}^0$ and $V_{\text{Er}}^{+}$ are the dominant states of the Er vacancy in Er$_2$O$_3$. In addition to the rather shallow thermodynamic transition level for the $(3−/0)$ state, we deduce that the Er vacancy is a potential shallow acceptor for producing p-type conduction, as it accepts electrons that are only 0.08 eV above the VBM and the $V_{\text{Er}}$ can be stabilized over the entire range of Fermi levels ($E_F$). From the thermodynamic transition level, we learn that the neutral state captures electrons with a thermal ionization barrier of only 0.08 eV $(3−/0)$ to form $V_{\text{Er}}^0$; considering the single-particle levels (TDOS), the $V_{\text{Er}}^0$ has a very shallow state near the VB edge, which implies high p-type conductivity near the VB edge. The $V_{\text{Er}}^0$ in Er$_2$O$_3$ has formation energies of 2.49 eV and 14.42 eV below the O-rich and Er-rich potential limits, respectively.

We have provided further calculations on the $V_{\text{Er}}$ in C-type Er$_2$O$_3$ during this revision with our limited computational resources. The neutral and +3 states of $V_{\text{Er}}$ ($V_{\text{Er}}^0$ and $V_{\text{Er}}^{+}$) are the defects in two typical charge states, of which the formation energies are 1.80 eV and 4.44 eV, respectively, for the C$_2$ site below the O-rich limit, and 1.97 eV and 4.66 eV, respectively, for the S$_8$ site below the O-rich limit. These display more stability compared with the A-type model owing to the fact that the formation enthalpy of C-type Er$_2$O$_3$ is lower than that of A-type Er$_2$O$_3$. A comparison of the formation energy of the defects shows that the $V_{\text{Er}}$ at the C$_2$ site is more stable than that at the S$_8$ site, with an energy difference of about 0.22 eV for $V_{\text{Er}}^0$ and 0.17 eV for $V_{\text{Er}}^{+}$, respectively. The thermodynamic transition levels for these defects for the charge transition between +3 and 0 lie at $E_V + 0.88$ eV for the C$_2$ site and $E_V + 0.90$ eV for the S$_8$ site, respectively, which are much deeper than the level in the A-type lattice, which is only about $E_V + 0.08$ eV (VBM, $E_V = 0$). Therefore, it is highly possible for such types of defect in the C-type lattice to induce a red-shift effect in the optical transitions that have been discussed. Moreover, it is very difficult to construct antiferromagnetic (AFM) Er$_2$O$_3$ in the C-type lattice, because it requires a unit cell that is more than twice as large as the 80-atom unit cell in C-type Er$_2$O$_3$. This will greatly increase the computational burden and, at the same time, AFM Er$_2$O$_3$ has not been observed in the C-type lattice.

Although it is necessary to discuss the C-type cubic structure of Er$_2$O$_3$ and discuss the differences in the distortions due to the cation vacancy defects, in this paper we initially focus on the A-type structure with the aim of providing a simple model of energy conversion, and we will provide a further detailed discussion in the following papers.

**Er interstitial (Er$_i$)**

An excess Er atom as an interstitial defect in Er$_2$O$_3$ may be a donor-like center that carries excess electrons. The neighboring O sites will be greatly distorted by such an Er$_i$ with hybridized bonding. The Er$_i$ will also form an Er–Er bond with neighboring Er sites. Fig. 6(a) shows the shallow donor-like eigenvalue for Er$_i^0$, which is 0.2 eV below the CBM, and its antibonding states are slightly higher than this occupied state with intervals of less than 0.43 eV. This indicates that the excess electrons of Er$_i^0$ occupy perturbed host states (PHS) below the CBM and rather than being localized in the gap. The shallow donor-like states of Er$_i^0$ lie 0.1–0.2 eV below the CBM as a PHS level. The distorted host Er site shifts the deep occupied 4f level to 0.6 eV higher than the VBM via a high-energy Er–Er bond, which is shown as the localized 4f orbital in Fig. 6(b). The doubly positive state (Er$_i^{2+}$) has only a 4f$^{12}$5s$^2$5p$^0$ shell, because the 6s$^2$ electrons have been ionized. The TDOS shows only a localized 4f state in the gap, which is about 0.6 eV higher than the VBM. In addition to that shown in Fig. 6(b), only one excess electron occupies the 4f orbitals from the nearest-neighbor Er site. The Er$_i^{2+}$ state shows that the electron that is stabilized in the 4f orbitals remains the same as in the case of Er$_i^{2+}$; the electron is localized in the $p$–$\pi$ orbital from the second nearest-neighbor O site instead of the 4f orbital of the first nearest-neighbor Er site, which shows the localization character of the 4f–4f electronic interactions among the Er sites.

The formation energies of Er$_i$ in different charge states within Er$_2$O$_3$ are shown in Fig. 6(c). The transition levels of both $(0+/1)$ and $(+/2+)$ states remain inside the CB, which shows that the electrons are easily ionized and resonant in the levels within the host CB. The transition level of the $(2+/3+)$ state is 1.80 eV above the VBM deep in the gap, which indicates a deep donor level in Er$_i^{2+}$, compared with the shallow donor behavior of Er$_i^0$. However, the formation energies of Er$_i^0$ are 10.35 eV and 22.28 eV below the Er-rich and O-rich limits, respectively. Therefore, this shallow donor state is unlikely to form for Er$_i^0$ and only deep donor states can be stabilized in the range of values of $E_F$ that are within the gap for Er$_i^{2+}$. The Er-related defect, especially in the form of an interstitial, is considered to be unlikely to have an influence on electron transport and is basically unstable due to its relatively high formation energy. Considering a similar discussion in a study by Chronoeos et al., the cation Frenkel defect is unlikely to occur due to the high defect reaction energies, which is based on a combination of a model of Born lattice potential and DFT calculations. Therefore, the Er$_i$ defect, as an interstitial, is also impossible to form.

**Anion Frenkel defect (a-Fr)**

The anion Frenkel (a-Fr) defect is a common pair defect that is mostly formed by thermal fluctuations, which lead to the escape of one anion from its original site to an interstitial site trapped in the lattice. Thus, it always consists of a vacancy and an interstitial with charge neutrality. In fact, there is another prominent pair defect called the Schottky (STK) defect in the crystal lattice, which consists of a pair vacancy of a cation and anion. However, the STK defect is energetically favored in the simple cubic structure as demonstrated in our previous work, whereas the a-Fr defect is more favorable in the CaF$_2$ (fluorite)-type structure. As supported by our defect formation energy, the neutral state of the STK defect (2$V_{\text{O}} + 2V_{\text{O}} + V_{\text{O}}$) (i.e., a unit of A-type Er$_2$O$_3$) exhibits a formation energy of as high as 7.16 eV in the Er$_2$O$_3$ host lattice. Actually, Er$_2$O$_3$ can be treated as a
Fig. 6  (a) TDOS of Er, in neutral (Er\(^0\)), singly positive (Er\(^+\)), doubly positive (Er\(^{2+}\)), and triply positive (Er\(^{3+}\)) charge states. The dashed line denotes the highest occupied level for electrons. (b) Localized electron and hole orbitals at the relaxed Er sites and nearby distorted Er and O sites (Er = green, O = red). (c) Formation energy of Er, below the O- and Er-rich chemical potential limits, displaying a "+U_{eff}" defect with a shallow donor feature between 0, +1 and +2 states.
distorted fluorite structure with broken symmetry; moreover, the radius of O is much smaller than that of Er. The a-Fr defect is a variable defect owing to the difference in crossover distance between the vacancy and the interstitial defect; furthermore, the coordination number of O required for creating the a-Fr defect is also different.

Fig. 7(a) shows the formation energies, which are dependent on the crossover distance, of three typical a-Fr defects (V\(_O\) + O\(_i\)) in Er\(_2\)O\(_3\). We found that for crossover distances between 3.3 Å and 4.2 Å there is a normal equilibrium state for most a-Fr defects, because the O\(_i\) is sufficiently far away to escape the Coulomb attractive potential generated by the V\(_O\) to prevent recombination and mutual annihilation. Another interesting state, which we called the metastable state, has a crossover distance of around 3.1–3.3 Å. The O\(_i\) remains at the cost of evident local site distortions to balance or screen the Coulomb attractive potential. From the formation energy, we see that the metastable O-related a-Fr defect is easy to form by external UV (~ 250 nm) photoirradiation with an excitation energy of 4–5 eV and is stabilized by lattice distortions induced by thermal fluctuations. From Fig. 7(a), we found that the a-Fr defect in the metastable state has the lowest formation energy of 4.01 eV in total, which means about 2.0 eV per defect site. The formation energy of a-Fr is actually independent of the chemical potentials but relates to the lattice distortions due to ambient thermal fluctuations.

Based on the formation energies of a-Fr within different charge states (−1, 0 and +1), we confirmed that a-Fr in both equilibrium and metastable states has a rather stable state at neutral charge in the mid-gap area. The a-Fr has a large positive value of \(U_{\text{eff}}\) (\(U_{\text{eff}} = +2.64\) eV), and the neutral state that is stabilized with respect to the \(E_F\) lies from \(E_V + 0.77\) eV (transition (0/+ for a-Fr) to \(E_V + 3.41\) eV (transition (−/0) for a-Fr) within the band gap area. Therefore, the reactions of O-related defects between various charge states are between deep levels that are generated by the neutral a-Fr.

Fig. 7(b) shows the single-particle levels (TDOS) of two typical states of the a-Fr defect in Er\(_2\)O\(_3\). The upper panel shows the TDOS of the a-Fr within the equilibrium spatial extent (> 3.3 Å). There are three localized p–p electronic states at the VB edge, which are similar to the states of O\(_i\) in the lattice. Another localized hole state lies 3.04 eV higher than the highest localized p–p level. In contrast, the metastable a-Fr only gives rise to a shallow p–p acceptor level close to the VBM, as shown in the lower panel of Fig. 7(b), and the hole state is 0.2 eV below the CBM. Fig. 7(c) shows the local orbitals of the gap states discussed above. The localized electronic states remain at the nearby distorted O sites of O\(_i\), whereas the localized hole orbital surrounds the space left by V\(_O\). We made a further comparison of the formation energy of the a-Fr pair defect vs. the crossover distance between Er\(_2\)O\(_3\) and CeO\(_2\), as shown in Fig. 7(d), as they have similar CaF\(_2\)-type local symmetry; we see that Er\(_2\)O\(_3\) exhibits a more drastic change in the distance of the a-Fr pair.

Pan et al. investigated the effects of oxygen vacancies in other lanthanide sesquioxides such as Lu\(_2\)O\(_3\).\(^{67}\) They show that both oxygen vacancies and interstitials (non-lattice defects) are
good charge carrier centers so as to enable filament conduction in \( \text{Lu}_2\text{O}_3 \).\(^{67}\) We similarly deduce that the oxygen-related defects play a similar role in \( \text{Er}_2\text{O}_3 \). This has gained support from the work of Chronoes et al.\(^{68}\) By their studies, it was found that all the sesquioxides favor the oxygen Frenkel pair defect (anion Frenkel, a-Fr), which is the lowest-energy mechanism of intrinsic lattice disorder in their calculations and summarizations. Stanek et al. have carried out further studies on the site preference of the sesquioxides and found that the lattice is less sensitive to such oxygen-related a-Fr pair defects.\(^{79}\) Therefore, in both the work of Stanek et al. and ours, these deductions are consistent with observations of the crystallography of these materials, which means that there is a high concentration of interstitial sites that exist in \( \text{O} \) sublattices, which are fully flexible to accommodate a large amount of oxygen defects.\(^{69}\) Moreover, a further experiment on an irradiation-induced order–disorder phase transformation in the sesquioxide \( \text{ Dy}_2\text{O}_3 \) by Stanek et al.\(^{71}\) shows that the lattice will become disorder under a high-energy irradiation process, so as to produce a large amount of anion disorder or deficient defects in the disordered fluorite structure, and most of these defects result from the oxygen sublattice. This experiment provides supportive information for our model, in that a high-intensity laser beam will produce a large number of a-Fr pair defects within the host lattice in both ordered and disordered fluorite structures. In the process towards the equilibrium state, the a-Fr pair defects will become mobile enough to form a possible superlattice,\(^{71}\) which means that a potential transporting effect is provided by the a-Fr with the help of lattice disorder or vibrations.

Therefore, we similarly consider that oxygen-related defects, in particular a-Fr pairs, are the key intrinsic defects that have subtle effects on the practical optical properties of the host materials. They have a great potential to generate spurious emissions or low efficiencies of emissions in phosphor applications, owing to their property of trapping electrons or holes at \( \text{O} \)-related defect sites. Accordingly, a-Fr pair defects may play the role of charge-trapping sites that trap charge carriers, either electrons or holes, to limit or influence the luminescence performance.

### Native defects-induced luminescence properties

An experiment reported persistent upconverted luminescence of \( \text{Er}_2\text{O}_3 \) under photostimulation at 975–980 nm \(( \sim 1.3 \text{ eV})\).\(^{14}\) A comparison was made between vacuum and ambient conditions, the only difference being the luminescence intensity. Persistent luminescence of \( \text{Er}_2\text{O}_3 \) has been discussed that was not related to extrinsic or intrinsic RE\(^{3+}\) ions. Therefore, the native point defects in this material play a significant role in modulating its luminescence properties. We summarize these native point defects in Table 1. Their thermodynamic transitions and single-particle levels are summarized in Fig. 8(a) and (b).

As we discussed above, there are many possible levels of native point defects that participate in or assist the process of luminescence. However, the key point for persistent luminescence is the release of energy at the activator site in the form of photons with energy transfer from the storage site. Therefore, we need to start with reactions of defect states for energy conversion during persistent luminescence. In contrast to other RE dopant-induced persistent luminescence in host materials, we introduce a model in which native point defects can be classified into two types, one acting as activators and the other being trap centers for energy storage. The most important point is to illustrate how electrons are driven by energy conversion for persistent luminescence. Here, we utilize the results of calculations on single-particle levels and thermodynamic transition levels (ionization energy) to elucidate our models.

For the process of persistent luminescence, we give an interpretation of the experimental results in the following way.

There are two possible mechanisms of luminescence, as illustrated in Fig. 8(c) and (d): one is related to \( F^- \) centers, the other is related to energy conversion between defects. It has been discussed that oxygen vacancies formed with \( +U_{\text{eff}} \) are usually color centers.\(^{24}\) \( V_{\text{O}6} \) and \( V_{\text{O}4} \) are the so-called \( F^- \) centers in \( \text{Er}_2\text{O}_3 \). The energy intervals between localized electronic and hole levels in the band gap are about 1.3 eV, by which electrons can be stimulated from the occupied levels in the band gap into the CB or near the CB edge via the localized hole states \((E_V + 4.46 \text{ eV for } V_{\text{O}6}, E_V + 4.93 \text{ eV for } V_{\text{O}4}, E_V = \text{VBM})\) in which the excited electrons are accommodated. The localized occupied levels for electrons that were mentioned lie at 3.30 eV and 3.67 eV for \( V_{\text{O}6} \) and \( V_{\text{O}4} \), respectively. An excited electron at the CB edge is relatively unstable compared with its defect-related states in the band gap; therefore, it will be de-excited back to its original ground-state levels at \( E_V + 3.30 \text{ eV for } V_{\text{O}4} \) and \( E_V + 3.67 \text{ eV for } V_{\text{O}4} \), respectively, and the energy released in the form of photons with energies of 2.21 eV \((561 \text{ nm})\) and 1.82 eV \((681 \text{ nm})\), respectively. The experimentally reported NIR emission is about 1540 nm owing to the de-excitation of electrons from the CB edge to the localized level for \( V_{\text{O}4} \). The energy difference between this state and the CB edge is about 0.7 eV \((1770 \text{ nm})\), which is of the correct magnitude compared with the experimental results.\(^{14}\) According to analysis based on the \( F^- \) center model (Fig. 8(c)), the experimentally reported twin peaks are attributed to the de-excitation of excited electrons back to the localized electron level within the band gap in the ground state for such defects. The lower luminescence band \((535 \text{ nm})\) seems to correspond to de-excitation at the \( V_{\text{O}6} \) site and the higher band \((660 \text{ nm})\) at the \( V_{\text{O}4} \) site, respectively.
The paths of the excitation and de-excitation of electrons are illustrated by the blue, pink and purple lines in Fig. 8(c).

However, this model seems to be inconsistent with the phenomenon observed in experiments. An experiment compared luminescence between vacuum and ambient conditions.\(^\text{14,72,73}\) The twin peak observed at 565 nm was found to be evidently weakened in ambient conditions compared with the peak found in a vacuum. We noted that the sample used in the experiment was a powder of nanoparticle size. As we know, nanoparticles have large surface-to-volume ratios; therefore, types of O vacancy dominate the entire luminescence of the particle. Oxygen has low coordination numbers on the upper surface owing to structural truncation compared with the bulk. On exposure to air, a surface with O vacancies with low coordination numbers is covered or terminated by H\(_2\)O molecules so as to quench the related luminescence. The other luminescence peak is attributed to O vacancies with high coordination numbers found in the bulk. A recent experiment on persistent luminescence shows a further comparison for the luminescence at 1540 nm between vacuum and ambient conditions; a similar weakening effect was found in ambient conditions, which means a reduction in the transition from V\(^{2+}\)O\(_4\) to V\(^{2+}\)O\(_6\).\(^\text{14}\) Because the peak position is unchanged, it is therefore a reduction in the levels induced by V\(_{06}\). Based on this discussion, the F\(^+\)-center model of persistent luminescence has been ruled out.

Another, different mechanism is a model of energy conversion relating to different defects that react mutually between different charge states so as to accomplish the energy transfer for luminescence. Therefore, in this model, we not only refer to the reactions of charge states in a given defect, but also discuss the reactions of inter-defects through different evolutions of charge states from the point of view of energy. In Fig. 8(d), the blue, pink and purple lines denote the possible paths of the excitation and de-excitation of electrons, which correspond to absorption and emission in luminescence spectroscopy. The two major differences from the F\(^+\)-center model are that the native point defects provide upconversion levels for electrons that are continuously stimulated from the VB into the localized defect levels in the gap and that, secondly, the luminescence is the release of energy in terms of photons among the reactions of different defects between various charge states, instead of single defect sites between different charge states.

Luminescence occurs with the transfer of electrons between the deep levels near the mid-gap area, as demonstrated by the following reactions of defects in the form of two-electron processes:

\[
\begin{align*}
2V_{06}^0 + O_i^0 & \rightarrow 2V_{06}^2 + O_i^2 + U_{\text{ef}1} \\
2V_{04}^0 + O_i^0 & \rightarrow 2V_{04}^2 + O_i^2 + U_{\text{ef}2} \\
2V_{06}^+ + O_i^0 & \rightarrow 2V_{06}^{2+} + O_i^2 + U_{\text{ef}3} \\
2V_{04}^+ + O_i^0 & \rightarrow 2V_{04}^{2+} + O_i^2 + U_{\text{ef}4} \\
O_i^{0+} + O_i^{0-} \rightarrow (a\cdot Fr)_{\text{metastable}} + U_{\text{ef}5}
\end{align*}
\]

(Taking into consideration the image charge corrections for charged states of point defects, the effective correlation...
energies for these related two-electron reactions of defects are $U_{\text{eff1}} = -3.91 \text{ eV}$, $U_{\text{eff2}} = -4.48 \text{ eV}$, $U_{\text{eff3}} = -1.77 \text{ eV}$, and $U_{\text{eff4}} = -3.31 \text{ eV}$, respectively, which show that the reactions are uniformly exothermic. The two electrons in these reactions will overcome the repulsive Coulomb potentials to become correlated to form a doubly occupied site, i.e. $O_i^–$, as a result of the negative $U_{\text{eff}}$. Hence, the energies that are released per electron from each of the defect reactions above are 1.95 eV, 2.24 eV, 0.88 eV, and 1.65 eV, respectively. We studied a luminescence experiment with peak values of 660 nm, 535 nm, 1540 nm, and 825 nm. Hence, our predicted energies can be shown as wavelengths if they are released in terms of photons: 635 nm, 554 nm, 1402 nm, and 749 nm with mean relative errors (MREs) of 4%, –3%, 9%, and 9%, respectively. Considering that an acceptable error is 50% from the experimental measurements made by Tanner et al.,14 we confirm that the predictions from our second model (the energy conversion model) show good consistency. According to this analysis, the wavelengths of the luminescence centers predicted by the energy conversion model not only exhibit the correct physical and chemical trends but also have relatively low errors. Based on discussions of the second model, we deduce that the short-wavelength limit of the emission spectra observed in the experiment on persistent luminescence14 is contributed by the energy converted by electron transfer from $V_{\text{O}6}$ to the $O_i$ sites. After eqn (2) and (3) complete the reactions, the final products, i.e. $V_{\text{O}6}^0$ and $O_i^{–}$, are converted back into the $a$-Fr in the metastable state, as the $a$-Fr is about 2.68 eV lower than the energy of $V_{\text{O}6}^0 + O_i^{–}$; therefore, $U_{\text{eff}} = -2.68 \text{ eV}$.

This can be confirmed from the single-particle levels as summarized in our TDOS calculations in Fig. 8(d). Electrons are excited by near-infrared (NIR) photostimulation with a photon energy of 975–980 nm (~1.3 eV). A many-photon process will assist the electrons to move from the valence band into the levels of the native point defects and then into the localized defect levels provided by $V_{\text{O}6}$ and $V_{\text{O}6}^0$, and the electrons that are originally from the valence band can be easily exchanged with the lowest occupied electron levels via thermal fluctuations within the range of 30–100 meV. The electrons will be further excited into the conduction band via a two-photon excitation process using NIR photostimulation at 980 nm (1.3 eV). As is known, Er$^{3+}$-related solids are widely employed in the field of optical communications owing to their unique emission at 1.54 μm.74 Compared with other La$^{3+}$ ion-related solids, Er$^{3+}$-related solids are preferred because of their highly quantum-efficient two-photon upconversion process.75 Actually, from Fig. 3(a) of ref. 14, the slope of the relationship between the power and the emission intensity shows that the process is a two-photon upconversion with a value of the slope of about 2.2 for measurements under vacuum conditions. Measurements in air show that about 0.8 photons are involved, which nearly approaches 1, and a one-photon process in air has also been confirmed from our transition levels.

These electrons are relatively unstable and will be de-excited spontaneously back to the closest localized defect levels provided by $V_{\text{O}6}^0$ and $V_{\text{O}6}^0$, as shown in Fig. 8(d). These levels are localized hole states and only accommodate these de-excited electrons for a short time (several ns); then, the electrons will finally be accepted/trapped by the deep localized levels provided by $O_i^–$ sites. The $O_i^–$ sites will be converted back into rather stable states in the form of $O_i^{2–}$ deep in the gap and transfer these electrons back to the occupied localized electron levels provided by $O_i^–$ sites. Afterwards, these electrons complete the entire local transport cycle back to the top of the VB because there are three localized electron levels for $O_i^{2–}$ sites with intervals of about 0.2–0.3 eV and the lowest localized electron levels are only about 0.15 eV higher than the top of the VB. To enable the existence of $O_i^{2–}$, it can be possible to create a number of $O_i$ sites locally in Er$O_3$ temporarily by optical excitation such as NIR photostimulation. From the optical aspect, with continuous excitation of the higher-energy absorption band, the electron paramagnetic resonance (EPR)-active spin-1/2 state $O_i$ further increases in concentration when temporally generated. The reactions shown in eqn (2) and (3) will cause quenching of this state via a rapid transition reaction of this defect, $2O_i^{–} \rightarrow O_i^{0} + O_i^{–}$, as it has a negative $U_{\text{eff}}$ as discussed in the section on O interstitials. Therefore, the $O_i$ is the essential activator site and the luminescent state is its EPR-active spin-1/2 state $O_i^{0}$, which is locally generated by temporary NIR photostimulation. With electron–hole recombination and the release of the energy in terms of photons, this luminescent state will revert to the $O_i^{–}$ state to transfer the electrons back to the VB of the host material.

According to our energy conversion model, we have made further efforts to generalize this model and explain the mechanism of persistent luminescence together with its photoelectron dynamics in terms of physical chemistry. Inspired by the radiation decay of the visible and invisible rays studied by Becquerel in 1896, we have transferred this mechanism of radiation decay to model our persistent luminescence, as they are similar to some extent: the emission of rays is the release of energy due to the transfer of electrons between different levels within an atom (radiation decay) or solid particles (persistent luminescence), respectively.

As shown in Fig. 9(a), we have here established a conversed chain reaction (CCR) scheme to reinforce our energy conversion model that is related to native defects, because we believe that this CCR scheme is a more general mechanism compared with the abovementioned reactions of native defects, as the native point defects are intermediates that assist in the operation of this mechanism. The reason why we have considered CCR is because the abovementioned reactions of native point defects in different charge states are exothermic processes that are accomplished within a short timescale that is similar to that of the lattice vibrations, and the final product of each cycle is the same as the initial reactant if we consider eqn (2)–(4).

Fig. 9(b) shows a schematic process of energy conversion. In each new chain reaction cycle, the concentration of the reactants is actually a reduced concentration of the products from the previous cycle. Therefore, we assume that there is a possibility $\eta$, which represents the efficiency of energy conversion per second, that a- Fr pair defects separate from each other into neutral states ($V_{\text{O}6}^0$ and $O_i^{0}$). We neglect the effect of temperature in the Arrhenius equation, the value of $\eta$ is nearly constant for a given material as an intrinsic property based on
Fig. 9  (a) Schematic diagram of the conversed chain reaction showing the energy conversions. (b) Detailed schematic process of energy conversion within a group of native point defects between different charge states. (c) A passive core–shell model is suggested to enhance the formation of a DAP band (DAP: donor–acceptor pair), so as to quench the emission at the short-wavelength limit for a red phosphor.
the synthesis conditions, and the entire reaction process together with energy conversion is confirmed as being based on a more complicated reaction model than a first-order chemical reaction. We mark each cycle based on the time in seconds. According to our modeling, the rate of the reaction may not be linearly dependent on the reactant concentration at that moment:

\[
A(t) = \left( 1 - \frac{1}{2} \eta \right)^{t} A_0
\]  
(5)

Only when \( \eta \) is a very small constant \( (0 < \eta \ll 1) \) do we get an approximation that is similar to the first-order reaction rate equation shown as follows:

\[
\frac{dA}{dt} \approx -\left( \frac{1}{2} \right)^{t} A_0
\]  
(5a)

Here, \( A(t) \) is the concentration of a-Fr defects with respect to the time after the external photoexcitation, and \( A_0 \) is the initial concentration. We see that eqn (5a) turns out to represent a first-order chemical reaction, which is similar to the phenomenon of radiation decay.

In an obvious relationship, the number of photons generated is proportional to the concentration of a-Fr defects at each moment in time. Then, regardless of the approximation from eqn (5) and (5a), the number of photons generated by the CCR at a time \( t \) is shown exactly as:

\[
N_{\text{photon}}(t) = \left( \frac{\eta}{2} \right)^{t} N_{\text{photon}}(0)
\]  
(6)

If we simply assume that the emission intensity is proportional to the number of photons, this allows \( I_{\text{PL}}(t) \propto N_{\text{photon}}(t) \). Thus, the spectrum of the decay of persistent luminescence can be simulated by an equation based on eqn (6):

\[
I_{\text{PL}}(t) = \left( \frac{\eta}{2} \right)^{t} I_{\text{PL}}(0) \quad \text{and}
\]

\[
t_{1/2} = \ln 2 / (\ln 2 - \ln \eta) \rightarrow \text{twin-peak PL}
\]  
(7a)

We find that the behavior of eqn (7a) is very similar to that in experiments on persistent luminescence.\(^{12-14,19}\) Eqn (5)–(7a) all represent special cases of luminescence with multiple peaks and between different charge states, which is due to O vacancies having the effect of a positive \( U_{\text{eff}} \). For persistent luminescence with a single peak and a \(-U_{\text{eff}}\) for O defects eqn (7a) is generalized as eqn (7b):

\[
I_{\text{PL}}(t) = (\eta)^{t}I_{\text{PL}}(0) \quad \text{and} \quad t_{1/2} = \ln 2 / (\ln 2 - \ln \eta) \rightarrow \text{single-peak PL}
\]  
(7b)

We see that the half-life is actually an intrinsic property of given materials that exhibit persistent luminescence. Considering both of the above cases, an increase in the value of \( \eta \) is the key factor that prolongs the duration of persistent luminescence. Otherwise, it is also possible to increase the initial intensity of luminescence by merely increasing the defect concentration, so as to increase the duration of persistent luminescence by rigidly shifting up the emission decay curve. It is highly technically demanding to increase the value of \( \eta \), which is essential for prolonging the intrinsic persistent luminescence of a given material. It is advisable to increase the defect concentration and reduce the activation barrier to defect reactions. Usually, sintering is a popular method for increasing the defect concentration by means of increasing the density of materials via a more densely packed structure, which has been reported in other experiments.\(^{76-79}\) Another possible way to create neutral pair defects in a state of thermal disequilibrium is by using external stress, which can provide a similar effect to sintering.\(^{80-84}\)

According to the experiment on the persistent luminescence of \( \text{Er}_2\text{O}_3 \) by Tanner et al.,\(^{14}\) we see that the decay of the emission at 660 nm has a half-life of about 0.25 s, whereas this is about 0.625 s for the decay at 1540 nm. Using eqn (7a), the values of \( \eta \) are found to be 0.125 (or 12.5%) for 660 nm and 0.660 (66.0%) for 1540 nm in vacuum conditions. This means that the rates of conversion from a-Fr to \( \text{V}_0^0 + \text{O}_i^0 \) are 12.5% and 14.4% in ambient air conditions (i.e. \( (1/2)\sqrt{0.660 \times 0.125} \)) for the processes shown in eqn (2) and (3). Our energy conversion modeling shows consistent results, in that the luminescence intensity is higher at 1540 nm than the case at 660 nm owing to the energy conversion efficiency being higher.

**Conclusion**

We have investigated the mechanism of the intrinsic persistent luminescence of \( \text{Er}_2\text{O}_3 \) in the A-type lattice based on first-principles calculations. Particular attention has been paid to the native point defects, thermodynamic transition levels and single-particle levels in terms of their electronic properties and formation energies. The native point defects are engaged in mutual subtle interactions in the form of chemical reactions between different charge states. The release of energy related to lattice distortion facilitates the conversion of energy for electrons to be transported between the valence band and the trap levels or even between the deep trap levels so as to generate persistent luminescence.

In the A-type lattice of \( \text{Er}_2\text{O}_3 \), we chose 4- and 6-fold coordinated \( \text{V}_0 \) to model this type of defect via low and high coordination numbers of \( \text{V}_0 \) in real materials. At the Er-rich limit (or O-poor chemical potential limit), we found that both \( \text{V}_0^0 \) and \( \text{V}_0^0 \) have nearly the same lowest energies with values of 0.13 eV and 0.05 eV, which indicates that high-coordination \( \text{V}_0 \) is energetically favored in \( \text{Er}_2\text{O}_3 \). Both are deep donors with a \(+U_{\text{eff}}\) and the charge state transitions and single-particle levels remain deep near the mid-gap region. At the O-rich limit, the neutral \( \text{O}_i \) is the dominant defect with a formation energy of as low as 2.16 eV and it has a \(-U_{\text{eff}} = -0.45 \) eV. The \( \text{V}_0^0 \) acts as a shallow acceptor with a transition level \((-3/2)\) that is only 0.08 eV higher than the VBM, and it has a relatively low formation energy of 2.49 eV in the same O-rich conditions with a rather high negative \(-U_{\text{eff}} \). For the lattice distortion defect a-Fr, we found two categories of state, equilibrium and metastable, which correspond to different ranges of the crossover distance for such pair defects. The metastable type has 4.01 eV per pair (i.e. 2.01 eV per defect site), whereas the value is 6.06 eV per pair.
for the equilibrium state. Er\textsubscript{i} in the triply positive state is relatively stable in Er\textsubscript{2}O\textsubscript{3} because it has a negative formation energy at the Er-rich limit, and its deep donor transition level (2\textsuperscript{+}/3\textsuperscript{+}) is 1.8 eV higher than the VBM. Except for V\textsubscript{Er}, nearly all of these native point defects have deep thermal transition levels that are related to E\textsubscript{T} near the mid-gap region.

An experiment has observed three different bands of persistent luminescence in the emission spectrum, which are located at 535 nm (green), 660 nm (red), and 1540 nm (NIR).\textsuperscript{14} We ruled out the F\textsuperscript{−}-center model of luminescence for persistent luminescence because comparative experiments on both ordinary upconversion luminescence and persistent phosphors between ambient air and vacuum conditions show quenching of the optical transition owing to the termination by water vapor of low-coordinate V\textsubscript{O} sites.\textsuperscript{14,72} A model of energy conversion has been established to explain the electron transport between different trap levels that supplies the energy to transfer and maintain the observed dynamics of persistent luminescence. Our calculations on the thermodynamic transition levels confirm that both visible and NIR experimentally observed intrinsic persistent luminescence (phosphor or afterglow) are related to the thermodynamic transition levels of O-related defects, and the thermodynamic transition levels within different charge states for these defects are independent of the chemical potentials of the given species. The lowest peak (short-wavelength limit) in the persistent luminescence band is contributed by the transition levels of charged donor-acceptor recombination between (0+/+) of V\textsubscript{O}\textsubscript{4} and (2−/0) of O\textsubscript{i}, in which the excited electrons and emitted photons exactly follow the zero-phonon lines, because the mutual local lattice relaxation effects of these pair defects counteract each other. In this way, similar physical and chemical trends show that the red luminescence (intermediate) is generated by the transitions between (0+/+) of V\textsubscript{O}\textsubscript{6} and (2−/0) of O\textsubscript{i}, whereas the persistent luminescence in the NIR region is provided by the levels between V\textsubscript{O}\textsubscript{4} and V\textsubscript{O}\textsubscript{6}, as summarized in Fig. 9(c).

We have further established a conversed chain reaction (CCR) model to interpret the energy conversion process of persistent luminescence in terms of the inter-reactions of native point defects between different charge states. It is advantageous to use the study of defect levels combined with that of formation energies to suggest limits to doping energy and explain photo-stimulated luminescence in terms of native point defects.

Acknowledgements

The author gratefully acknowledges the support of the Natural Science Foundation of China (NSFC) for the Youth Scientist grant (Grant No. NSFC 11504309) and the initial start-up grant support from the Department General Research Fund (Dept. GRF) from ABCT in the Hong Kong Polytechnic University.

References


4 J. Hölsä, Electrochim. Soc. Interface, 2009, 18, 42.


18 B. Huang, Philos. Mag., 2014, 94, 3052.


