Designing Semiconductor Heterostructures Using Digitally Accessible Electronic-Structure Data

March 31, 2016

Semiconductor sandwich structures, so-called heterojunctions, are at the heart of modern applications with tremendous societal impact: Light-emitting diodes shape the future of lighting and solar cells are promising for renewable energy. They use heterojunctions for optimal efficiency and desirable performance. However, their computer-based design is hampered by the high cost of atomistic techniques used to select materials based on alignment of valence and conduction bands. We describe, validate, and demonstrate a technique that addresses this grand challenge: It provides fast access to band alignment, allowing rapid screening of existing online databases to find suitable, previously manufactured materials for applications. We predict 11,506 possible LEDs and 117 possible solar cells using our broadly applicable, automated method.

1 Background

The lack of a fast, reliable approach for purposeful selection of materials for semiconductor heterojunction design currently hampers progress for widely used applications with tremendous societal impact, such as solid-state lighting and photovoltaic energy production. The heterojunction in light-emitting diodes (LEDs) and photovoltaic cells is similar: An active component—the emitter in an LED or the absorber in a solar cell—is determined by application-driven criteria, while electron- and hole-transport layer materials can be picked more freely. Good candidates are lattice-matched with the active material, but, more importantly, show suitable energy alignment of valence- and conduction bands. Being able to quickly and reliably screen large numbers of materials for band alignment is an *unsolved* grand challenge. Achieving this would lead to higher efficiency of semiconductor LEDs, transition novel photovoltaic absorbers into actual devices more rapidly, and guide the search for photocatalytic water splitters or 2D electron gases.

Publicly accessible online databases have compiled both experimental and computational data for numerous materials properties of tens to hundreds of thousands of materials. In particular, the relatively low computational cost of density functional theory (DFT) [1] recently lead to projects that contain both structural information and detailed electronic-structure data such as band structures and densities of states [2–5] for materials that were synthesized before and are reported in the International Crystal Structure Database. The DFT-based databases have previously been used to design scintillator materials [6], to study optical effects in calcites [7], to find trends in total energy and enthalpy of formation [8], for Li-ion battery design [9], to screen electrocatalytic materials [10], and to identify novel solar-cell absorbers [11]. At the same time, there are very few studies [12, 13] that explicitly use electronic-structure data available in these databases for design of functional materials.

This situation can partly be attributed to the notorious band-gap problem of DFT, which detrimentally affects band structures stored in DFT-based databases. Since the *dispersion* of Kohn-Sham eigenvalues computed within DFT oftentimes is in better agreement with experiment, we believe that the wealth of digitally accessible electronic-structure data in *currently existing* online databases can be explored to either directly find suitable materials for semiconductor heterostructures, or to at least narrow the vast search space of possible candidates.

Screening for lattice-matching can be carried out straightforwardly because data on atomic geometries and lattice parameters is readily available [2–5]. However, information on band alignment generally does not exist since it requires much higher computational effort. In this report we describe, validate, and demonstrate an approach to obtain band alignment and offsets at the interface between different materials, exclusively using electronic-structure data that is currently digitally stored in online databases. We validate our results with respect to the experimental and computational literature and we demonstrate our technique by proposing electron- and hole-transport layers for heterojunction LEDs with CdSe and InP as emitters and for a solar-cell with an organo-metal halide ($CH_3NH_3PbI_3$) absorber. Based on these initial results, we show that our approach addresses the long-standing problem of materials design for semiconductor heterojunctions [14–19]. Pending experimental validation, we envision that it will enable the broad scientific community to design specific semiconductor heterojunctions at rapid pace in the future.

2 Methods

At the interface between two semiconductors, where bulk band structures of both materials merge into each other, a "transition region" forms: Band bending due to different doping profiles typically extends significantly into the material, while band-edge discontinuities (i.e. valence-band and conduction-band offsets of two materials with different band gaps) are confined to not more than a few atomic layers (see e.g. Ref. [20]). The relative energy position of valence-band maxima and conduction-band minima at the interface is called "band alignment" and decides whether the interface acts as barrier or a conductor for electrons or holes.

Atomistic first-principles calculations have been applied to quantify band alignment at semiconductor interfaces: It has been traced back to the absolute energy position of atomic core levels [21], to the electronic transition level of hydrogen impurities [22], or to the alignment of the vacuum levels for the materials in contact [23]. While being successful and accurate, the disadvantage of these techniques is their high computational cost. Large super cells with tens or even hundreds of atoms (either including defects/impurities or material slabs and vacuum) and quasiparticle energies are needed [18]. Since this is expensive even for specific individual interfaces using DFT, the information needed for those alignment approaches is not included in currently existing online databases, rendering a quick screening of thousands of material combinations impossible.

Such a rapid screening is, however, desirable for heterojunction design: Just the one online database (out of many) that we are using for this report, Materials Project [2, 24], contains band structures for 42,465 materials to date. For three-component heterostructures, where no two components are identical and the active material is in contact with exactly one hole-transport and one electron-transport layer, this would produce ${}^{42{,}465}P_3 \approx 7.8 \times 10^{13}$ possible arrangements. Even if the active material is fixed, a total of $\approx 1.8 \times 10^9$ possibilities remain.

In the literature an approach exists [15, 18, 25] that uses the "branch-point energy" (BPE) E_{BP} (also "charge neutrality level" or "effective mid-gap energy") as a universal energy level for band alignment, assuming negligible interface dipoles [20]. The BPE can be entirely traced back to the bulk band structure of a given semiconductor [15]. This is advantageous for our purposes, since this approximation neglects complicated and oftentimes unknown structural details of the interface itself. Initially, the BPE was described using energies at isolated $\mathbf{k}$-points [14] or as a Brillouin zone (BZ) average based on the first of Baldereschi's special points [26]. Later, a sum over a full Monkhorst-Pack [27] $\mathbf{k}$-point grid was used to compute E_{BP} according to [25]

$$E_{\mathrm{BP}} = \frac{1}{2N_{\mathbf{k}}} \sum_{\mathbf{k}} \left(\frac{1}{N_c} \sum_{i}^{N_c} \varepsilon_{c_i}^{\mathrm{QP}}(\mathbf{k}) + \frac{1}{N_v} \sum_{i}^{N_v} \varepsilon_{v_i}^{\mathrm{QP}}(\mathbf{k}) \right). \tag{1}$$

In Eq. (1) $N_{\mathbf{k}}$ is the number of points in the $\mathbf{k}$ mesh and the energies $\varepsilon_{c_i}^{\mathrm{QP}}(\mathbf{k})$ and $\varepsilon_{v_i}^{\mathrm{QP}}(\mathbf{k})$ describe the quasiparticle band structure of a material. N_c and N_v are the number of conduction and valence bands, respectively, included in the sum. Previously those were computed from the number of s and p valence electrons, N_e, which lead to an uncertainty on the order of 0.2 eV for E_{BP} [25]. In general, the numbers of bands N_c and N_v included in the average of Eq. (1) are parameters of this approach.

Here, we optimized N_c and N_v as free parameters by calculating the mean absolute error between the calculated and experimental E_{BP} for the 21 semiconductors listed in Table 1 where experimental

data is available. Using $N_c = N_e/8$ and $N_v = N_e/4$, a mean absolute difference of 0.22 eV from the experimental values in Table 1 was found. Halving the number of bands used produced a mean absolute difference of 0.31 eV, while doubling the number of bands led to a mean absolute difference of 0.73 eV. Further testing examined using equal number of valence and conduction bands: Setting $N_c = N_v = N_e/8$, the mean absolute difference was 0.33 eV and for $N_c = N_v = N_e/4$, the mean absolute difference was 0.67 eV. For this reason we used $N_c = N_e/8$ and $N_v = N_e/4$ here. The number of non-d and non-f valence electrons, N_e, is calculated using the Pymatgen `periodic_table` and `full_electronic_structure` modules. From these we obtain the total number of electrons with principle quantum number equal to the maximum principle quantum number for each species and sum those.

Here we will show that it is sufficient to sum only over path segments connecting two high-symmetry points (HSP) in the full Brillouin zone (FBZ), as provided by Materials Project, instead of taking each **k**-point of a Monkhorst-Pack grid into account in Eq. (1). Weighing each segment j correctly according to its symmetry-dependent multiplicity n_j and length L_j leads to

$$E_{\mathrm{BP}} = \frac{1}{2L_{\mathrm{path}}} \sum_{j}^{N_{\mathrm{seg}}} n_j L_j \sum_{\mathbf{k}\in j} \frac{1}{N_{\mathbf{k},j}} \left(\frac{1}{N_c} \sum_{i}^{N_c} \varepsilon_{c_i}^{\mathrm{QP}}(\mathbf{k}) + \frac{1}{N_v} \sum_{i}^{N_v} \varepsilon_{v_i}^{\mathrm{QP}}(\mathbf{k}) \right). \tag{2}$$

In Eq. (2) N_{seg} is the total number of segments connecting beginning and end of the band-structure plot (e.g. $\mathrm{X}-\mathrm{L}-\Gamma-\mathrm{M}$ corresponds to N_{seg}=3), n_j is the multiplicity of segment j, and $N_{\mathbf{k},j}$ is the number of **k**-points in the j^{th} segment. L_{path} is the total length of the path, i.e., the sum of lengths of all segments, $L_{\mathrm{path}} = \sum_j^{N_{\mathrm{seg}}} L_j n_j$. Since Materials Project does not provide information on multiplicities, we use n_j=1 for all segments. As we show in detail below, this does not introduce a significant error to our results, compared to the overall error bar of the method. Hence, we can now compute E_{BP} exclusively using electronic-structure data available in online databases.

Finally, DFT is not well-suited to provide accurate quasiparticle energies for Eq. (2). Instead, we use the linear correction scheme of Ref. [6] to rigidly open the gap of band structures provided by Materials Project with a scissor operator. The shift is derived from a linear least-squares fit between DFT and experimental gap for a large number of materials. Currently, our implementation can also use experimental band gaps directly, if available for a given material. In the future we are planning to explore the machine-learning scheme of Ref. [28] that uses the Kohn-Sham band gap along with other information from DFT results as predictor for the true band gap of a material.

E_{BP} computed via Eq. (2) allows us to *very quickly* align the bands of any material for which Materials Project provides a band structure plot. However, we need to verify two approximations: First, we need to check how large of an error is introduced when the sum over the full BZ is replaced by a sum over the band structure using n_j=1. Second, E_{BP} according to Eq. (2) neglects any interface-specific details and we need to validate that it is consistent with experimental results. In the following section, we discuss these two aspects of our approach based on the mean absolute errors of BPEs for common semiconductors. We then combine our scheme with "filters" and apply it to the problem of two heterojunction LEDs and a photovoltaic absorber material.

3 Results and Discussion

We first validate our approach of using the band-structure **k**-point path provided by Materials Project to compute E_{BP} according to Eq. (2), using equal multiplicities. For comparison, FBZ band structure calculations are performed using the Vienna *Ab-initio* Simulation Package [29, 30]. In Fig. 1 we compare the difference of E_{BP} from our results to a calculation with FBZ sampling for 30 common semiconductors. We find that the absolute error is not larger than 0.4 eV for any of the studied materials. The same is true when comparing the result for the band-structure **k**-point path with correct multiplicities to the one where each segment is weighed equally (n_j=1): At least 25 of the 30 E_{BP} even agree to within 0.2 eV. These results show that our approach is capable of providing E_{BP} consistent with results from FBZ sampling, however it has the advantage of extreme speed and straightforward applicability to *any* existing, digitally available band-structure data.

Next, we validate against experiment: For this purpose, Fig. 2 shows the difference between E_{BP} obtained using FBZ sampling as well as our technique with respect to experimental data for 21 semiconductors. Comparing Fig. 2(a) and (b) shows that both distributions look very similar and are centered around ≈ 0.2 eV, with the largest outlier showing -0.8 eV deviation. 17 of the 21 materials show an absolute deviation of ≤ 0.4 eV compared to experiment [see Fig. 2(b)]. While this is very satisfactory given the extreme speed and simplicity of our technique, the agreement is not perfect and this error bar needs to be kept in mind for applications [18].

Our validation results, based on the data in Table 1, are summarized in Fig. 3: In this plot the calculated E_{BP} for all 30 trial semiconductors are compared to literature values. The mean absolute difference between our results and experimental values for E_{BP} is 0.22 eV. The mean absolute difference between experimental values and the hydrogen-level alignment of Ref. [22] is 0.31 eV. Comparing experiment and the data in Ref. [18], the mean absolute difference is 0.18 eV, which is comparable to our method, but requires computationally much more expensive $GW\Gamma^1$@HSE calculations. We conclude that our technique is promising and well-suited for a fast initial screening of the band alignment for very large numbers of materials.

Finally, we provide a proof of concept by applying our technique to possible three-component heterostructures, two LEDs and a solar cell. Band structures were queried sequentially by Materials Project ID using the Pymatgen Open-Source Library [31]. The valence band maximum (VBM), conduction band minimum (CBM), energy gap (E_{g}), and E_{BP} for each compound were determined. Semiconductor band structures are gapped, i.e. $E_{\mathrm{CBM}} - E_{\mathrm{VBM}} > 0$ eV, however, since DFT underestimates band gaps, especially small-gap semiconductors can appear gapless in DFT. To avoid excluding these, material data was retained for further screening if $E_{\mathrm{CBM}} - E_{\mathrm{VBM}} > -1.0$ eV. Figure 3 shows six examples (Ge, CdO, GaSb, InAs, InSb and InN) of semiconductors which are predicted to be gapless by Materials Project, but become gapped upon application of the linear correction. At this stage, information was gathered for a total of 37,727 materials.

Next, the linear gap correction [6] was applied, which shifts E_{BP} according to Eq. (2) by half the correction. All materials that are not solid under ambient conditions (e.g. mp-20066 corresponds to CO_2) were filtered out using the density stored in Materials Project, only keeping those with $\rho > 2$ g/cm^3. Unstable materials were removed by retaining only those with energies above hull (E_{Hull}) that are smaller than 0.01 eV. Finally, compounds which contain radioactive (Ac, Th, U, Np, Pu, Tc, Pm) or very expensive elements (Pt, Pd) were filtered out as well. In the remaining candidate list, there were 2,347 binary compounds or 10,405 materials if ternaries were included.

At this point, heterojunctions were proposed for two LED emitter materials (CdSe and InP) and one promising novel solar cell absorber ($CH_3NH_3PbI_3$). After fixing these as active component of the heterojunction, we used the band alignment criteria shown in Fig. 4 to find possible electron and hole transport layers. This entire process is visualized diagrammatically in Fig. 5 and has been implemented as iPython notebook. The flexible iPython implementation is compatible with the Pymatgen library and allows users to easily alter, add, or remove any of the filters, and select the number of materials and the band offsets between materials. At the same time, down-selecting a maximum possible number of 1.8×10^9 heterostructures using this (not optimized) code did not take longer than 10 hours on a single CPU.

The number of heterostructures which pass all filters and satisfy the band alignment criteria of Fig. 4 are given in Table 2. Several potentially viable compositions are read off and listed in Table 3. Among these, Zn_2GeO_4 has been experimentally identified as an efficient electron transport layer for electroluminescent devices [32]. Other experimentally tested electron transport layers such as TiO_2 [33] and ZrO_2 [34] are not present. Both of these materials were excluded by our tool because their gaps exceeded the upper cutoff that we imposed for the LED structure. In the case of TiO_2, the calculated gap is 4.52 eV compared to an experimental gap of 3.2 eV. The linear gap correction overcorrects in this case. Shifting the VBM and CBM such that the calculated gap matches the experimental gap, TiO_2 is still excluded because after aligning E_{BP} the CBM is 1.7 eV higher than the one in CdSe and 1.2 eV above the CBM of InP. For ZrO_2, the corrected band gap is 5.60 eV, close to the experimental gap of 5.7 eV. Further, the CBM is more than 2.0 eV over the CBMs of CdSe and InP. Literature data is not consistent on the issue of relative positions of bands. The specific values for band offsets in our framework will be left to the users' discretion.

Among the hole transport layers listed in Table 3 WO_3 and V_2O_5 have both been experimentally found to be suitable [35]. Other viable hole transport layers [35] are omitted due to DFT band gap errors, predicting CuO and Co_3O_4 to be gapless. If the gaps are adjusted to the experimental values of 1.35 eV for CuO [36] and 1.6 eV for Co_3O_4 [37] the materials are still excluded. For a material to be a transport layer according to our criteria, it must have a gap at least 0.9 eV larger than the emitting layer, however, the calculated gaps are 1.60 eV (CdSe) and 1.47 eV (InP). MoO_3 [34] is excluded from our list of electron transport layers due to its CBM being 0.13 eV below the CBM of CdSe, 0.03 eV more than allowed by the alignment criteria we used. NiO [38] is excluded due to its VBM being 1.24 eV above the VBM of CdSe and 0.59 eV above the VBM of InP. These examples do not indicate a failure of our general procedure, but hint at a more careful specification of the band alignment criteria. Those are not fully fixed in the literature either: In experiment cases were reported where CBM of the electron-transport layer is below [39] or above [40] the CBM of the emitter. At the same time, these criteria strongly influence the total number of predicted heterostructures. For the present report, we did not yet choose optimal criteria, but this is planned for the near future.

4 Conclusions and Outlook

In this work we describe, validate, and demonstrate a technique that addresses the grand challenge of achieving fast and purposeful computational design of band alignment in semiconductor heterostructures. Band offsets were determined by aligning the branch-point energies of materials, as computed directly and exclusively from band-structure data accessible via the Materials Project database. To demonstrate the applicability of this technique, currently available digital

data for over 42,000 previously fabricated materials was analyzed and filtered to produce human-readable lists of potentially realizable three-component LEDs and solar-cell heterostructures. A maximum possible number of 1.8×10^9 heterostructures for each case was down-selected to about 100 – 10,000 possibilities within no more than 10 hours on a single CPU.

As the branch-point energy is a bulk property, interface dependence or influence of doping is not included. While this has to be kept in mind when assessing the error bars of this technique, it is also the reason for the extremely low computational cost. At the same time, the type of data provided for a given material by currently existing online databases dictates what can be used for a more detailed analysis. Systematic improvement, e.g. for the most promising heterostructure candidates, can be achieved, at significantly higher computational cost, using full DFT or many-body perturbation theory simulations. Forthcoming experimental work will seek to measure performance of heterostructure compositions predicted in this work.

While presently only data from Materials Project is taken into account, the extension of our technique to any other database that provides direct access to band-structure data is straightforward and planned, for instance for OQMD and Aflowlib, in the near future. Furthermore, the code developed in this work is implemented as iPython notebooks, which makes it readily applicable for other users. Parameters such as the number of materials and the desired band offsets between materials can easily be altered. Filters can be added, modified, or removed as needed to search for desired heterostructure configurations and band offsets. These iPython notebooks will be made publicly available in the near future as well.

We envision that our technique will be broadly applied, either for direct design of desired semiconductor heterostructures, or to reduce the vast candidate search space. This can tremendously accelerate the search for novel high-efficiency LEDs, solar cell devices, photocatalytic water splitters, or 2D electron gases confined to semiconductor interfaces and, hence, contribute to applied research with broad societal impact as well as addressing fundamental scientific problems.

Acknowledgments

The authors are grateful to M. Shim for helpful discussions regarding fabrication considerations, to S. Biesboer, B. Cabinian, A. Erwin, and H. Wu for the first version of the Python code that this study builds on, and J. Leveillee for providing band structure data of $CH_3NH_3PbI_3$. This work was supported in part by the Materials and Manufacturing Graduate Student Fellowship of the National Center for Supercomputing Applications. This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Part of this research is supported by the National Science Foundation (awards: CBET-1437230 and DMR-1555153).

Availability of Supporting Data

All digital data used for screening for electron- and hole-transport layer materials in this work was obtained from Materials Project [2, 24] using the Pymatgen Open-Source Library [31]. The band structure data of the two LED emitter materials studied here, CdSe and InP, were also taken from Materials Project. A realistic use case is demonstrated based on band structure data of $CH_3NH_3PbI_3$ that is not currently available on Materials Project. Since the active component is not part of the materials screening carried out here, the only information needed for this material is E_g = 1.86 eV and E_{BP} = 1.28 eV above the valence band maximum.

References

[1] W. Kohn and L. J. Sham: Self-Consistent Equations Including Exchange and Correlation Effects. In: *Phys. Rev.* **140**, A1133–A1138 (1965), DOI: 10.1103/PhysRev.140.A1133.

[2] A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder, and K. A. Persson: The Materials Project: A materials genome approach to accelerating materials innovation. In: *APL Materials* **1** (1), p. 011002 (2013), DOI: 10.1063/1.4812323.

[3] R. H. Taylor, F. Rose, C. Toher, O. Levy, K. Yang, M. B. Nardelli, and S. Curtarolo: A RESTful API for exchanging materials data in the AFLOWLIB.org consortium. In: *Comp. Mater. Sci.* **93**, pp. 178–192 (2014), DOI: 10.1016/j.commatsci.2014.05.014.

[4] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton: Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). In: *JOM* **65** (11), pp. 1501–1509 (2013), DOI: 10.1007/s11837-013-0755-4.

[5] C. Ortiz, O. Eriksson, and M. Klintenberg: Data mining and accelerated electronic structure theory as a tool in the search for new functional materials. In: *Comput. Mater. Sci.* **44** (4), pp. 1042–1049 (2009), DOI: 10.1016/j.commatsci.2008.07.016.

[6] W. Setyawan, R. M. Gaume, S. Lam, R. S. Feigelson, and S. Curtarolo: High-Throughput Combinatorial Database of Electronic Band Structures for Inorganic Scintillator Materials. In: *ACS Comb. Sci.* **13** (4), pp. 382–390 (2011), DOI: 10.1021/co200012w.

[7] K. M. Poduska, L. Regev, E. Boaretto, L. Addadi, S. Weiner, L. Kronik, and S. Curtarolo: Decoupling Local Disorder and Optical Effects in Infrared Spectra: Differentiating Between Calcites with Different Origins. In: *Advanced Materials* **23** (4), pp. 550–554 (2011), DOI: 10.1002/adma.201003890.

[8] A. M. Deml, R. O'Hayre, C. Wolverton, and V. Stevanović: Predicting density functional theory total energies and enthalpies of formation of metal-nonmetal compounds by linear regression. In: *Phys. Rev. B* **93**, p. 085142 (2016), DOI: 10.1103/PhysRevB.93.085142.

[9] M. N. Obrovac and V. L. Chevrier: Alloy Negative Electrodes for Li-Ion Batteries. In: *Chem. Rev.* **114** (23), pp. 11444–11502 (2014), DOI: 10.1021/cr500207g.

[10] J. Greeley, T. Jaramillo, J. Bonde, I. Chorkendorff, and J. Nørskov: Computational High-Throughput Screening of Electrocatalytic Materials for Hydrogen Evolution. In: *Nat. Mater.* **5** (11), pp. 909–913 (2006).

[11] L. Yu and A. Zunger: Identification of Potential Photovoltaic Absorbers Based on First-Principles Spectroscopic Screening of Materials. In: *Phys. Rev. Lett.* **108**, p. 068701 (2012), DOI: 10.1103/PhysRevLett.108.068701.

[12] I. E. Castelli, F. Hüser, M. Pandey, H. Li, K. S. Thygesen, B. Seger, A. Jain, K. A. Persson, G. Ceder, and K. W. Jacobsen: New Light-Harvesting Materials Using Accurate and Efficient Bandgap Calculations. In: *Adv. Ener. Mater.* **5** (2), p. 1400915 (2015), DOI: 10.1002/aenm.201400915.

[13] S. Korbel, M. A. L. Marques, and S. Botti: Stability and electronic properties of new inorganic perovskites from high-throughput ab initio calculations. In: *J. Mater. Chem. C* (2016), DOI: 10.1039/C5TC04172D.

[14] F. Flores and C. Tejedor: Energy barriers and interface states at heterojunctions. In: *J. Phys. C Solid State* **12**, pp. 731–749 (1979), DOI: 10.1088/0022-3719/12/4/018.

[15] J. Tersoff: Theory of semiconductor heterojunctions: The role of quantum dipoles. In: *Phys. Rev. B* **30**, pp. 4874–4877 (1984), DOI: 10.1103/PhysRevB.30.4874.

[16] A. Schleife, C. Rödl, F. Fuchs, J. Furthmüller, and F. Bechstedt: Optical and energy-loss spectra of MgO, ZnO, and CdO from ab initio many-body calculations. In: *Phys. Rev. B* **80**, p. 035112 (2009), DOI: 10.1103/PhysRevB.80.035112.

[17] M. Stengel, P. Aguado-Puente, N. A. Spaldin, and J. Junquera: Band alignment at metal / ferroelectric interfaces: Insights and artifacts from first principles. In: *Phys. Rev. B* **83**, p. 235112 (2011), DOI: 10.1103/PhysRevB.83.235112.

[18] Y. Hinuma, A. Grüneis, G. Kresse, and F. Oba: Band alignment of semiconductors from density-functional theory and many-body perturbation theory. In: *Phys. Rev. B* **90**, p. 155405 (2014), DOI: 10.1103/PhysRevB.90.155405.

[19] C. G. Van de Walle and R. M. Martin: Theoretical calculations of heterojunction discontinuities in the Si/Ge system. In: *Phys. Rev. B* **34**, pp. 5621–5634 (1986), DOI: 10.1103/PhysRevB.34.5621.

[20] W. Mönch: *Semiconductor Surfaces and Interfaces*. Springer, Berlin 2001.

[21] S.-H. Wei and A. Zunger: Calculated natural band offsets of all II-VI and III-V semiconductors: Chemical trends and the role of cation *d* orbitals. In: *Appl. Phys. Lett.* **72** (16), pp. 2011–2013 (1998), DOI: 10.1063/1.121249.

[22] C. G. Van de Walle and J. Neugebauer: Universal alignment of hydrogen levels in semiconductors, insulators and solutions. In: *Nature* **423**, pp. 626–628 (2003), DOI: 10.1038/nature01665.

[23] R. Anderson: Experiments on Ge-GaAs heterojunctions. In: *Solid State Electron.* **5** (5), pp. 341–351 (1962), DOI: 10.1016/0038-1101(62)90115-6.

[24] A. Jain, G. Hautier, C. J. Moore, S. P. Ong, C. C. Fischer, T. Mueller, K. A. Persson, and G. Ceder: A high-throughput infrastructure for density functional theory calculations. In: *Comput. Mater. Sci.* **50** (8), pp. 2295–2310 (2011), DOI: 10.1016/j.commatsci.2011.02.023.

[25] A. Schleife, F. Fuchs, C. Rödl, J. Furthmüller, and F. Bechstedt: Branch-point energies and band discontinuities of III-nitrides and III-/II-oxides from quasiparticle band-structure calculations. In: *Appl. Phys. Lett.* **94**, p. 012104 (2009), DOI: 10.1063/1.3059569.

[26] M. Cardona and N. E. Christensen: Acoustic deformation potentials and heterostructure band offsets in semiconductors. In: *Phys. Rev. B* **35**, pp. 6182–6194 (1987), DOI: 10.1103/PhysRevB.35.6182.

[27] H. J. Monkhorst and J. D. Pack: Special points for Brillouin-zone integrations. In: *Phys. Rev. B* **13**, pp. 5188–5192 (1976), DOI: 10.1103/PhysRevB.13.5188.

[28] J. Lee, A. Seko, K. Shitara, K. Nakayama, and I. Tanaka: Prediction model of band gap for inorganic compounds by combination of density functional theory calculations and machine learning techniques. In: *Phys. Rev. B* **93**, p. 115104 (2016), DOI: 10.1103/PhysRevB.93.115104.

[29] G. Kresse and J. Furthmüller: Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. In: *Comp. Mater. Sci.* **6**, pp. 15–50 (1996), DOI: 10.1016/0927-0256(96)00008-0.

[30] G. Kresse and D. Joubert: From ultrasoft pseudopotentials to the projector augmented-wave method. In: *Phys. Rev. B* **59**, pp. 1758–1775 (1999), DOI: 10.1103/PhysRevB.59.1758.

[31] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder: Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. In: *Comput. Mater. Sci.* **68**, pp. 314–319 (2013), DOI: 10.1016/j.commatsci.2012.10.028.

[32] J. Wang, C. Yan, S. Magdassi, and P. S. Lee: Zn_2GeO_4 Nanowires As Efficient Electron Injection Material for Electroluminescent Devices. In: *ACS Applied Materials & Interfaces* **5** (15), pp. 6793–6796 (2013), DOI: 10.1021/am401234a.

[33] N.-G. Park, J. van de Lagemaat, and A. J. Frank: Comparison of Dye-Sensitized Rutile- and Anatase-Based TiO2 Solar Cells. In: *J. Phys. Chem. B* **104** (38), pp. 8989–8994 (2000), DOI: 10.1021/jp9943651.

[34] N. Tokmoldin, N. Griffiths, D. D. C. Bradley, and S. A. Haque: A Hybrid Inorganic-Organic Semiconductor Light-Emitting Diode Using ZrO_2 as an Electron-Injection Layer. In: *Adv. Mater.* **21** (34), pp. 3475–3478 (2009), DOI: 10.1002/adma.200802594.

[35] Z. B. Wang, M. G. Helander, M. T. Greiner, J. Qiu, and Z. H. Lu: Analysis of charge-injection characteristics at electrode-organic interfaces: Case study of transition-metal oxides. In: *Phys. Rev. B* **80**, p. 235325 (2009), DOI: 10.1103/PhysRevB.80.235325.

[36] F. P. Koffyberg and F. A. Benko: A photoelectrochemical determination of the position of the conduction and valence band edges of *p*-type CuO. In: *J. Appl. Phys.* **53** (2), pp. 1173–1177 (1982), DOI: 10.1063/1.330567.

[37] J. van Elp, J. L. Wieland, H. Eskes, P. Kuiper, G. A. Sawatzky, F. M. F. de Groot, and T. S. Turner: Electronic structure of CoO, Li-doped CoO, and $LiCoO_2$. In: *Phys. Rev. B* **44**, pp. 6090–6103 (1991), DOI: 10.1103/PhysRevB.44.6090.

[38] J.-M. Caruge, J. E. Halpert, V. Bulović, and M. G. Bawendi: NiO as an Inorganic Hole-Transporting Layer in Quantum-Dot Light-Emitting Devices. In: *Nano Lett.* **6** (12), pp. 2991–2994 (2006), DOI: 10.1021/nl0623208.

[39] A. K. Chandiran, M. Abdi-Jalebi, M. K. Nazeeruddin, and M. Grätzel: Analysis of Electron Transfer Properties of ZnO and TiO_2 Photoanodes for Dye-Sensitized Solar Cells. In: *ACS Nano* **8** (3), pp. 2261–2268 (2014), DOI: 10.1021/nn405535j.

[40] J. M. Caruge, J. E. Halpert, V. Wood, V. Bulovic, and M. G. Bawendi: Colloidal quantum-dot light-emitting diodes with metal-oxide charge transport layers. In: *Nat. Photon.* **2** (4), pp. 247–250 (2008), DOI: 10.1038/nphoton.2008.34.

[41] W. Mönch: *Electronic properties of semiconductor interfaces*. Springer, Berlin 2004.

[42] A. Dimoulas, P. Tsipas, A. Sotiropoulos, and E. K. Evangelou: Fermi-level pinning and charge neutrality level in germanium. In: *Appl. Phys. Lett.* **89** (25) (2006), DOI: 10.1063/1.2410241.

[43] L. Piper, P. Jefferson, T. Veal, C. McConville, J. Zúñiga-Pérez, and V. Muñoz-Sanjosé: X-ray photoemission studies of the electronic structure of single-crystalline CdO(100). In: *Superlattices Microst.* **42**, pp. 197–200 (2007), DOI: 10.1016/j.spmi.2007.04.029.

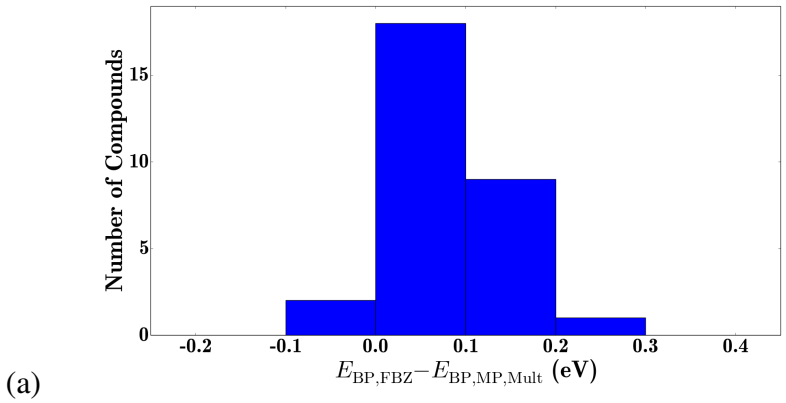
[44] D. Mourad, J.-P. Richters, L. Gérard, R. André, J. Bleuse, and H. Mariette: Determination of valence-band offset at cubic CdSe/ZnTe type-II heterojunctions: A combined experimental and theoretical approach. In: *Phys. Rev. B* **86**, p. 195308 (2012), DOI: 10.1103/PhysRevB.86.195308.

[45] Intrinsic Electron Accumulation at Clean InN Surfaces. In: *Phys. Rev. Lett.* **92**, p. 036804 (2004), DOI: 10.1103/PhysRevLett.92.036804.

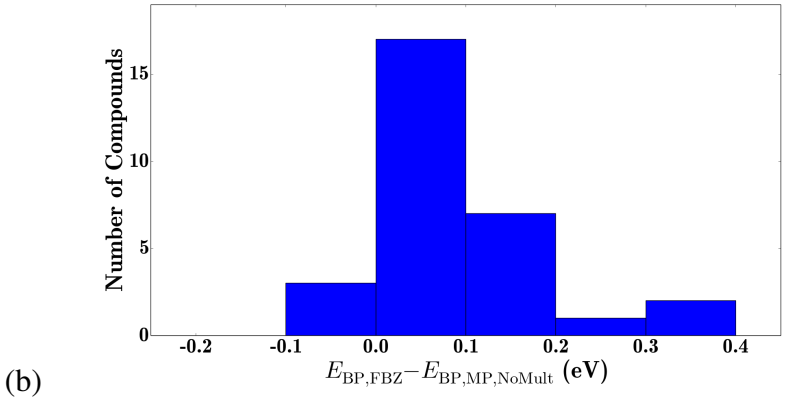
[46] P. D. C. King, T. D. Veal, P. H. Jefferson, J. Zúñiga-Pérez, V. Muñoz-Sanjosé, and C. F. McConville: Unification of the electrical behavior of defects, impurities, and surface states in semiconductors: Virtual gap states in CdO. In: *Phys. Rev. B* **79**, p. 035203 (2009), DOI: 10.1103/PhysRevB.79.035203.

[47] S. Kasap and P. Capper: *Springer Handbook of Electronic and Photonic Materials*. 1st, Springer US, 2007, ISBN: 978-0-387-26059-4.

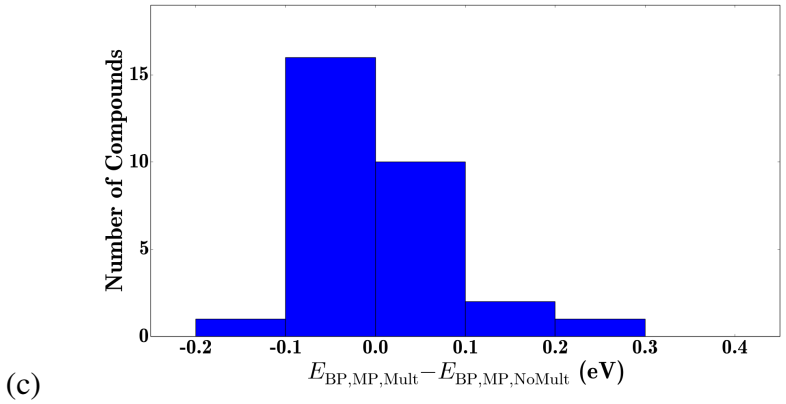
[48] N. B. Guy LeLay Jacques Derrien: *Semiconductor Interfaces: Formation and Properties*. 1st, Springer-Verlag, 1987, ISBN: 3642729673.



(a)



(b)



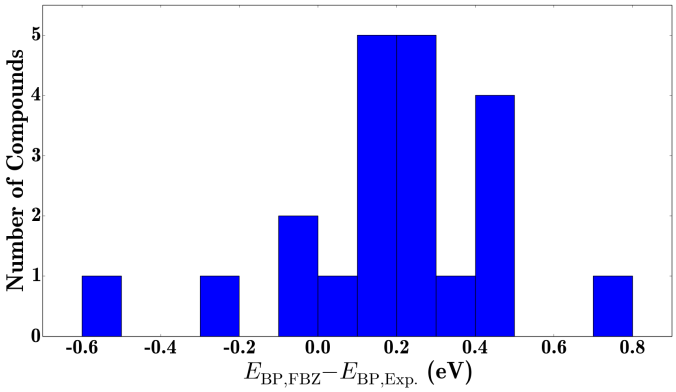
(c)

Figure 1: Difference of E_{BP} between (a) FBZ sampling and band-structure **k**-path sampling with correct segment multiplicities, (b) FBZ sampling and band-structure **k**-path sampling with equal segment multiplicities (n_j=1), and (c) band-structure **k**-path sampling with correct and with equal segment multiplicities. The explicit numerical data is given in Table 1.

Material	$E_{\mathrm{BP,MP}}$	$E_{\mathrm{g,MP}}$	$E_{\mathrm{BP,FBZ}}$	$E_{\mathrm{BP,Exp.}}$	$E_{\mathrm{BP,Lit}}$ [22]	$E_{\mathrm{BP,Lit}}$ [18]
Si	0.56	1.74	0.47	0.36 [41]	0.34	0.56
Ge	0.06	0.91	0.08	0.09 [42]	−0.25	0.01
CdO	1.64	0.24	2.03	1.3±0.1 [43]	–	–
KBr	4.09	6.72	4.27	–	–	–
MgO	4.84	6.90	5.10	–	–	–
AlAs	1.28	2.94	1.34	0.92 [41]	1.05	1.04
AlP	1.56	3.11	1.56	1.13 [41]	1.74	1.49
AlSb	0.91	2.57	0.93	0.53 [41]	0.08	0.19
CdSe	1.47	1.60	1.62	1.83 [44]	1.24	1.63
CdTe	1.24	1.71	1.34	–	0.61	0.75
GaAs	0.58	1.17	0.67	0.52 [41]	0.46	0.45
zb-GaN	2.05	3.03	2.40	–	–	–
GaP	1.06	3.06	1.11	0.83 [41]	1.01	0.93
GaSb	0.28	0.91	0.34	0.16 [41]	−0.25	0.33
InAs	0.63	0.91	0.73	0.50 [41]	0.19	0.38
InP	0.89	1.54	0.98	0.86 [41]	0.65	0.82
InSb	0.45	0.91	0.51	0.22 [41]	−0.34	0.25
ZnS	2.04	3.63	2.18	2.05 [41]	2.04	2.23
ZnSe	1.59	2.49	1.73	1.48 [41]	1.45	1.75
ZnTe	1.22	2.37	1.29	1.09 [44]	0.50	0.71
HfO_2	3.09	6.33	3.11	2.62 [41]	–	–
AlN	3.51	6.38	3.55	–	2.90	–
CdS	1.88	2.42	1.93	–	1.70	1.97
wz-GaN	2.32	3.26	2.4	2.4 [41]	–	–
InN	1.51	0.91	1.55	1.64 ± 0.1 [45]	1.91	–
ZnO	2.68	1.90	2.68	3.2 [46]	3.49	3.01
Al_2O_3	5.71	8.80	5.70	–	–	–
$LaAlO_3$	2.84	6.33	2.87	–	–	–
PbI_2	2.24	4.18	2.31	–	–	–
SiO_2	5.21	8.60	5.28	4.9 [41]	4.37	–

Table 1: Explicit numerical data for E_{BP} (in eV) using the different approaches discussed in the text for all 30 trial semiconductors. All values of E_{BP} are given relative to the valence band maxima.

(a)



(b)

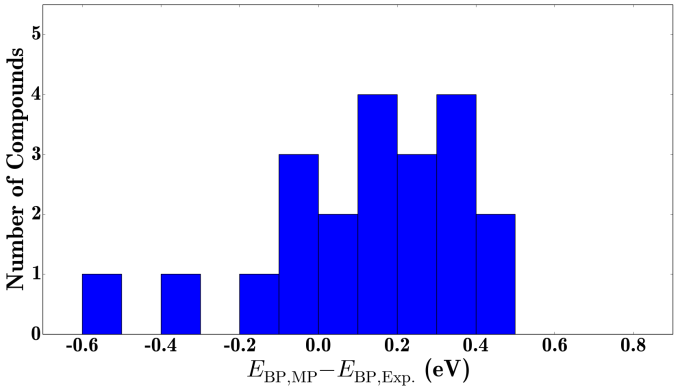


Figure 2: Difference of E_{BP} between experiment and (a) FBZ sampling, (b) band-structure **k**-path sampling with equal segment multiplicities. The explicit numerical data is given in Table 1.

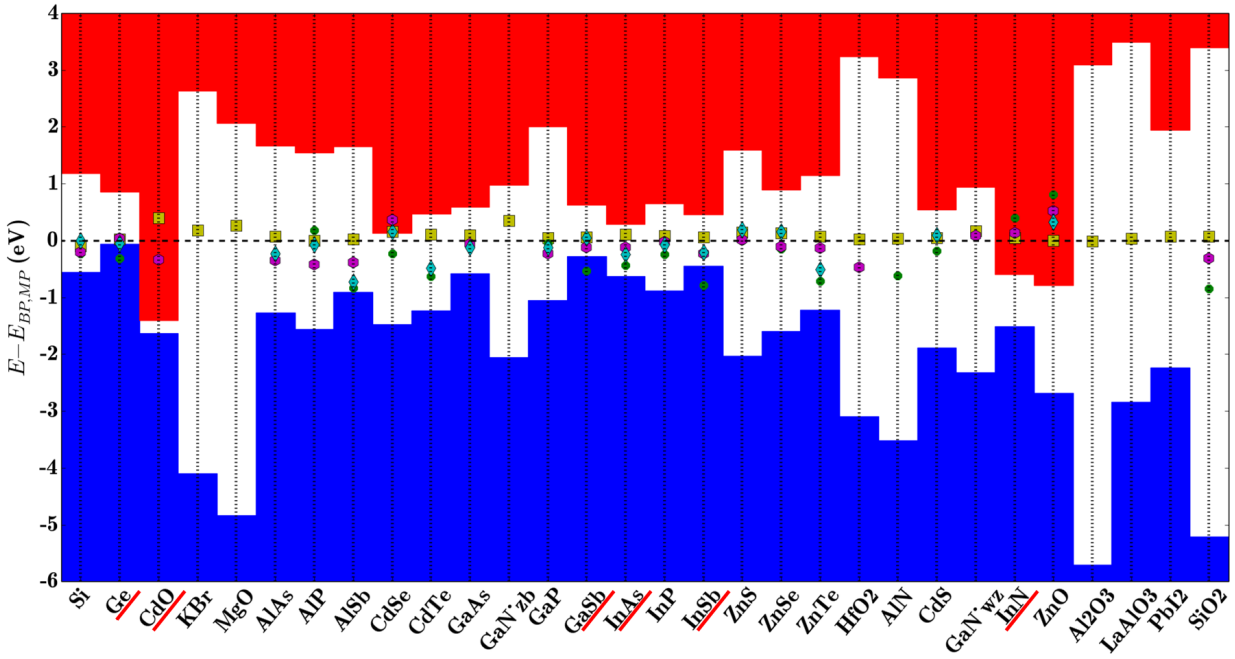


Figure 3: Band alignment for a set of 30 common semiconductors: E_{BP}, computed for the high-symmetry path from Materials Project using Eq. (2), is used as energy zero. Blue (red) bars indicate valence (conduction) bands, the white space in between is the band gap. Materials underlined in red are predicted to be gap-less within DFT, prior to application of the gap correction scheme [6]. Yellow squares are results obtained from FBZ sampling, green dots are computational results from Ref. [22], cyan diamonds are computational results from Ref. [18], violet hexagons are experimental results from Refs. [42, 43, 47, 48], see also Table 1. E_{BP} was calculated from Figure 2 of Ref. [18] by selecting the energy 0.56 eV above the Si VBM as the E_{BP}, matching our calculated result for Si. We assumed E_{BP} were aligned at the same energy (using the vacuum level in Ref. [18]) for all materials and calculated the position relative to the VBM.

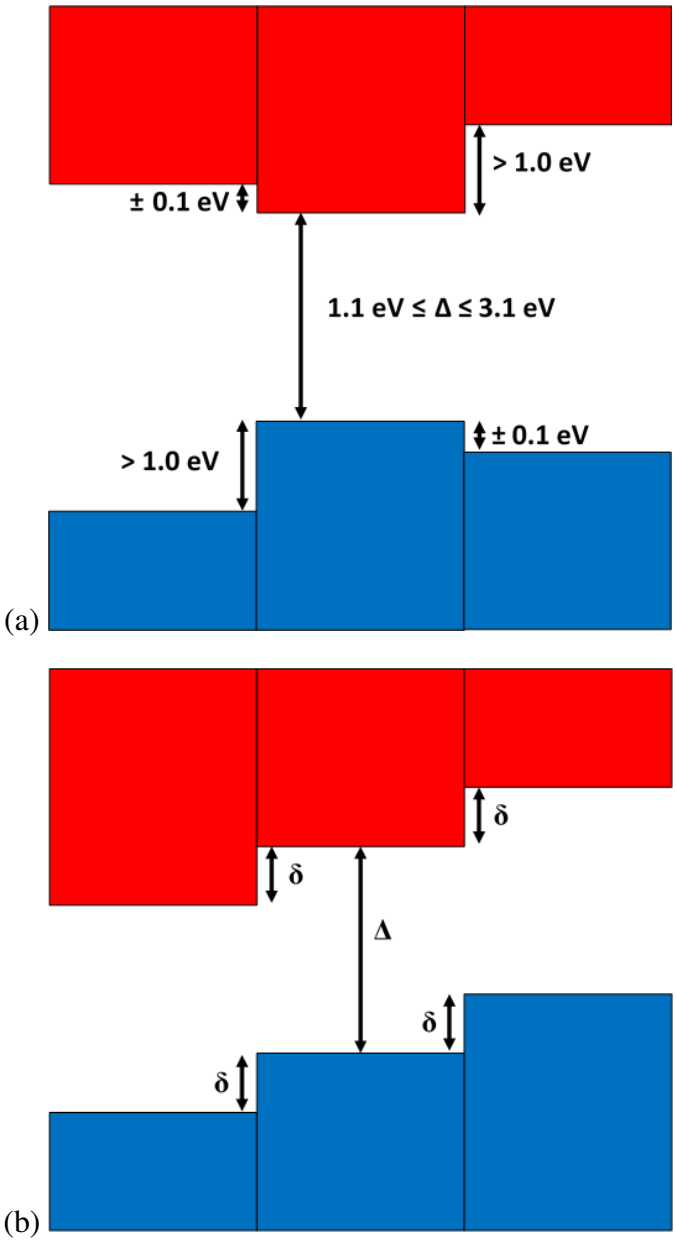


Figure 4: Band structure configurations assumed in this work for (a) LED and (b) solar cell. Blue (red) bars indicate valence (conduction) bands, the white space in between is the band gap. For the solar cells, $\delta = 0.5 \pm 0.1$ eV. A further requirement of 1.1 eV $\leq E_{\mathrm{g}} \leq$ 4.0 eV is imposed on the electron and hole transport layers.

Heterostructure	Binary	Ternary
LED-CdSe Center	92	9415
LED-InP Center	0	2091
Solar Cell-$CH_3NH_3PbI_3$	0	117

Table 2: Number of possible three-component heterostructures that satisfy all criteria applied for our demonstration use cases.

Heterostructure	Electron Transport Layers	Hole Transport Layers
LED-CdSe Center	As_2O_5, Zn_2GeO_4	AlP, WO_3, V_2O_5
LED-InP Center	$AlCuO_2$, $ZnSeO_4$	GaP, $MoSe_2$
Solar Cell-$CH_3NH_3PbI_3$	Ca_6GaN_5	$FeSi_2$, GeTe

Table 3: Promising examples picked from the lists of possible electron- and hole-transport layer materials for our demonstration use cases.

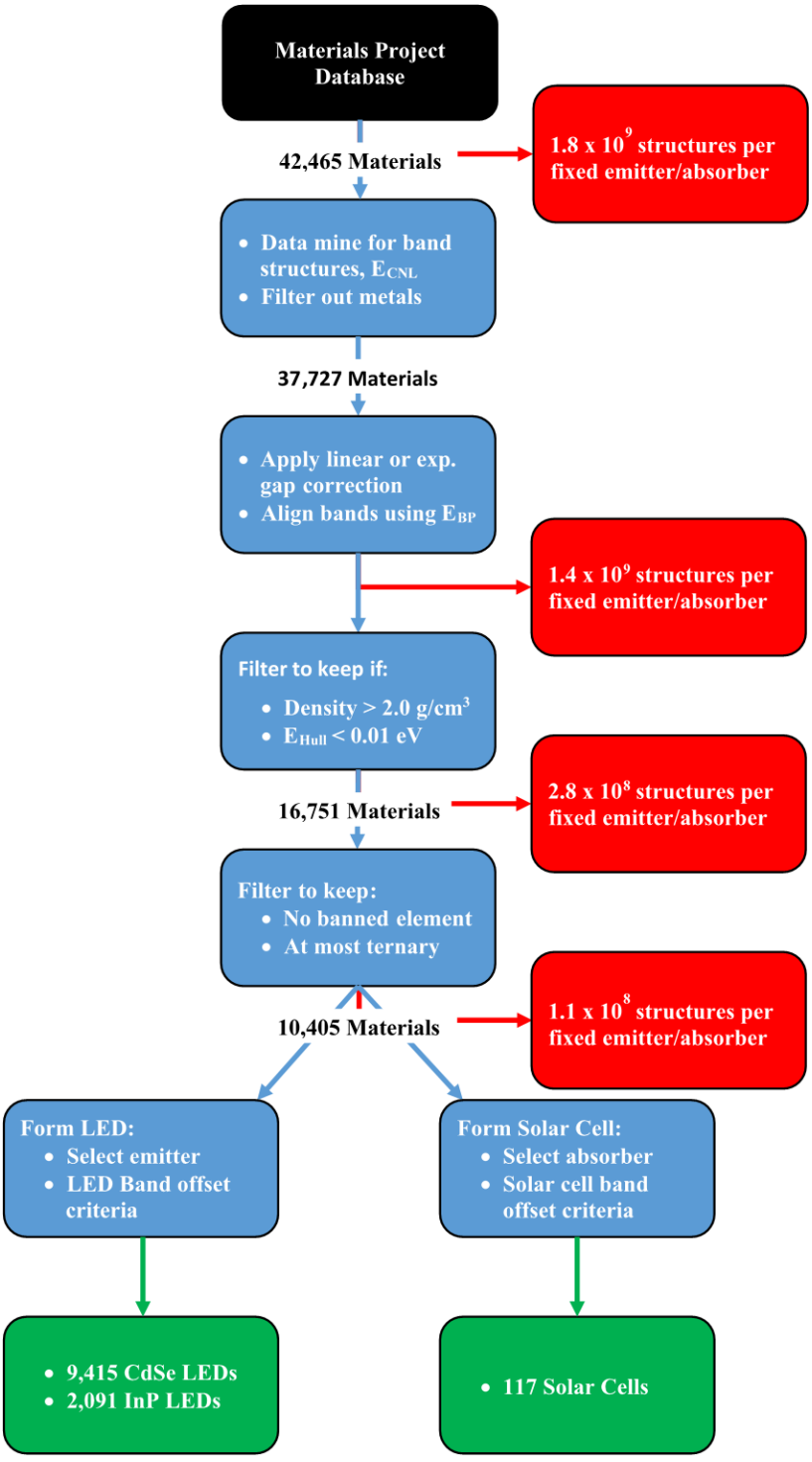


Figure 5: Flow chart used to filter materials for LEDs and solar cells. Blue boxes indicate steps in which either materials data is collected or modified. Red boxes indicate the number of heterostructures remaining after each level of filtering. Green boxes at the bottom are the final lists of heterostructures.