The strongest upper bounds on the axion mass come from astrophysical observations like the neutrino burst duration of SN1987A, which depends on the axion couplings to nucleons, or the white-dwarf cooling rates and red-giant evolution, which involve the axion-electron coupling. It has been recently argued that in variants of Dine-Fischler-Srednicki-Zhitnitsky (DFSZ) models with generation-dependent Peccei-Quinn (PQ) charges an approximate axion-nucleon decoupling can occur, strongly relaxing the SN1987A bound. However, as in standard DFSZ models, the axion remains in general coupled to electrons, unless an ad hoc cancellation is engineered. Here we show that axion-electron decoupling can be implemented without extra tunings in DFSZ-like models with three Higgs doublets. Remarkably, the numerical value of the quark mass ratio $\mu/m_d \approx 1/2$ is crucial to open up this possibility.

I. INTRODUCTION

It has been recently argued [1] that in variants of Dine-Fischler-Srednicki-Zhitnitsky (DFSZ) [2,3] models with two Higgs doublets and generation-dependent Peccei Quinn (PQ) charges, it is possible to strongly suppress axion couplings to nucleons (axion nucleophobia). This implies that the upper limit on the DFSZ axion mass from the neutrino burst duration of the supernova (SN) SN1987A, which is particularly strong and generally considered ineludible, can in fact be sizeably relaxed. The parameter space region that opens up for nucleophobic axions is, however, only marginal. This is because in DFSZ models the axion also couples to electrons, and then limits from anomalous cooling of white-dwarfs and red-giants, which are only moderately less restrictive, apply. On the other hand, generation-dependent PQ charges imply that the axion couplings to the mass eigenstate fermions receive corrections from intergenerational mixing. In Ref. [1] this type of effect was invoked to arrange for a tuned cancellation between two contributions to the axion-electron coupling: one proportional to the electron PQ charge, and the other coming from intergenerational mixing between the leptons. This allows the construction of models of nucleophobic and electrophobic axions that can evade all the tightest astrophysical bounds (astrophobic axion models). Although the tuning of the cancellation required to achieve a significant level of electrophobia is at the level of 10%, astrophobic axion models constructed in this way are admittedly not particularly elegant. In this short paper we put forth a more natural way to achieve astrophobia, which requires extending the scalar sector by a third Higgs doublet, but does not involve any ad hoc cancellation between different contributions to the axion-electron coupling. Remarkably, this mechanism implies a strong correlation between the couplings whereby the higher the level of suppression of the axion-nucleon coupling, the more the axion becomes electrophobic. Intriguingly, the mechanism can be implemented thanks to the fact that the light quark mass ratio is close to $m_u/m_d \approx 1/2$. 

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI. Funded by SCOAP3.
II. THE CONDITIONS FOR NUCLEOPHOBIA

Let us first recall the conditions for nucleophobia. We define the axion couplings to nucleons via

\[ \frac{C_N}{2f_a} \partial_\mu a \gamma^\mu r_s N, \]

with \( N = p, n \), while the fundamental couplings of the axion to quarks \( C_q \), with \( q = u, d, \ldots \), are defined from a similar expression by replacing \( N \to q \) and \( C_N \to C_q \). \( C_N \) can be expressed in terms of \( C_q \) using nonperturbative inputs from nucleon matrix elements [9]. To understand the mechanism behind axion-nucleon decoupling, it is convenient to consider the two linear combinations \( C_p \pm C_n \) and express them in terms of the \( C_q \). This yields

\[ C_p + C_n = 0.50(C_u + C_d - 1) - 2\delta_s, \]

\[ C_p - C_n = 1.27(C_u - C_d - f_{ud}), \]

where, in the second line, \( f_{ud} = f_u - f_d \) with \( f_{ud} = m_{ud}/(m_d + m_u) \) the model-independent contributions induced by the axion coupling to gluons, chosen in such a way that the axion does not mix with the \( N \). In the first line, \( l = f_u + f_d \) is an exact number, while \( \delta_s = 0.038C_u + 0.012C_d + 0.009C_b + 0.0035C_s \) [9] is a small \( \mathcal{O}(5\%) \) correction dominated by the \( s \)-quark contribution. Nucleophobia requires \( C_p \pm C_n \approx 0 \), which is possible in variant DFSZ models with two Higgs doublets \( H_{1,2} \) and nonuniversal PQ charge assignment [1]. To see this, let us focus on the first generation Yukawa terms

\[ \tilde{q}_1u_1H_1 + \tilde{q}_1d_1\tilde{H}_2, \]

where \( \tilde{H}_2 = i\sigma_2H_2^* \). The axion couplings to the light quark fields (neglecting flavor mixing, which is assumed to be small throughout this paper\(^2\)) are

\[ C_u = \frac{1}{2N}(X_{u_1} - X_{q_1}) = \frac{X_1}{2N}, \]

\[ C_d = \frac{1}{2N}(X_{d_1} - X_{q_1}) = \frac{X_2}{2N}. \]

Here \( X_{u_1} = X(u_1) \), etc. denote the PQ charges of the fermion fields while \( X_{1,2} = X(H_{1,2}) \). The coefficient of the PQ color anomaly then is

\[ 2N = \sum_{i=1}^{3}(X_{u_i} + X_{d_i} - 2X_{q_i}). \]

It is also convenient to define the contribution to the color anomaly from light quarks only:

\[ 2N_e = X_{u_1} + X_{d_1} - 2X_{q_1} = X_2 - X_1. \]

The first condition for ensuring approximate nucleophobia then reads [cf. Eq. (2)]

\[ C_u + C_d = \frac{N_e}{N} = 1, \]

i.e., only models in which the color anomaly is determined solely by the light \( u, d \) quarks (while the contributions from the two heavier generations cancel or vanish identically) have a chance to be nucleophobic.\(^3\) As emphasized in [1], this implies that nucleophobic axion models can be realized only if the PQ charges are generation-dependent. Assuming that the first condition Eq. (9) is satisfied, the second condition [cf. Eq. (3)] reads

\[ C_u - C_d = \frac{-X_1 - X_2}{2N} = \frac{X_1 + X_2}{X_1 - X_2} = f_{ud}. \]

Finally, by imposing the condition which ensures that the physical axion field is orthogonal to the Goldstone field of hypercharge \( U(1)_Y \), i.e., \( X_1v_1^2 + X_2v_2^2 = 0 \), we obtain a relation between the ratio of the vacuum expectation values (VEVs) \( v_{1,2} = \langle H_{1,2} \rangle \) and the ratio of the quark masses that must be satisfied in order to ensure nucleophobia:

\[ \frac{v_2^2}{v_1^2} = \frac{X_1}{X_2} = \frac{m_d}{m_u}. \]

With only two Higgs doublets responsible for breaking the electroweak symmetry and for providing masses to all the fermions, the lepton sector is unavoidably charged under the PQ symmetry and, as mentioned in the Introduction, electrophobia can only be enforced by tuning a cancellation between the contribution to the axion electron coupling proportional to the electron PQ charge, and corrections arising from lepton flavor mixing [1]. A possible, and more elegant alternative, is to introduce a third Higgs doublet \( H_3 \) with PQ charge \( X_3 \) that couples only to the leptons, and verify if the condition \( X_3 \approx 0 \) can be consistently implemented. In this case the whole lepton sector would be approximately neutral under the PQ symmetry, and in

\(^2\)In the presence of flavor mixing, \( C_q \to C_q + \Delta C_q \), where \( \Delta C_q \) involves quark mass diagonalization matrices. We refer the reader to [1] for details.

\(^3\)It is worthwhile mentioning that a certain number of models sharing precisely this property were found in a recent study of \( U(1) \) flavor symmetry for the quark sector [10].
particular the axion would decouple from the electrons. This possibility is explored in the remainder of the paper.

III. THE CONDITIONS FOR ELECTROPHOBIA

Let us consider a three-Higgs doublet model (3HDM) wherein $H_{1,2}$ couple to quarks as above, while $H_3$ couples to the leptons. We want to study if electrophobia can be implemented consistently with nucleophobia. Assuming Eq. (9) is satisfied, we have four additional conditions: orthogonality between the physical axion and the hypercharge Goldstone, the second condition for nucleophobia Eq. (11), and two conditions on the PQ charges that follow from the requirement that the four $U(1)$ rephasing symmetries of the kinetic term of the four scalar fields $H_{1,2,3}$ and $\phi^i$ (the latter being the Standard Model (SM) singlet, with $\chi_q = 1$, responsible for PQ breaking) are broken down to $U(1)_Y \times U(1)_{PQ}$. These last two conditions can be implemented either by coupling the leptonic Higgs doublet $H_3$ to both hadronic Higgses ($H_{1,2}$), or by coupling one of the two hadronic Higgses to the other two doublets:

$$H_3^* H_1 \phi^m + H_2^* H_2 \phi^n \text{ or } H_3^* H_{1,2} \phi^m + H_2^* H_1 \phi^n.$$  

(13)

For renormalizable operators one has, without loss of generality, $m = 1, 2$ and $n = \pm 1, \pm 2$, where negative values of $n$ mean Hermitian conjugation $\phi^n \equiv (\phi^i)^{\dagger n}$. All in all, for the first case in Eq. (13) the four conditions read

$$\frac{\chi_1}{\chi_2} = -\frac{m_d}{m_u},$$

(14)

$$\chi_1 v_1^2 + \chi_2 v_2^2 + \chi_3 v_3^2 = 0,$$

(15)

$$-\chi_3 + \chi_1 + m = 0,$$

(16)

$$-\chi_3 + \chi_2 + n = 0.$$  

(17)

To see if there is a consistent charge assignment that allows us to decouple the axion from the leptons, let us set $\chi_3 \to 0$. In this limit, Eq. (15) reduces to the previous condition, Eq. (12), while Eqs. (16)–(17) imply $\chi_1 / \chi_2 = m / n$ which, together with Eq. (14), yields

$$\frac{m_d}{m_u} = -\frac{m}{n}.$$  

(18)

Hence, with the first choice of operators in Eq. (13), electrophobia can be consistently implemented for the following values of the light quark mass ratio: $m_u / m_d = 2, 1, 1/2$. It is a fortunate coincidence that the actual value $m_u / m_d = 0.48(3)$ [9] is perfectly compatible with the first possibility. This renders it possible to have electrophobia together with nucleophobia by means of a suitable assignment of PQ charges, rather than by tuning of some parameters. By contrast, if the breaking $U(1)^4 \to U(1)_Y \times U(1)_{PQ}$ is enforced via the second set of operators in Eq. (13), respectively with $H_1$ or $H_2$ in the first term, electrophobia would require $m_d / m_u = 1$ or $\infty$ in the first case, and $m_d / m_u = 1$ or 0 in the latter. Hence for both these cases electrophobia would not be compatible with nucleophobia.

IV. ASTROPHOBIC AXIONS IN A 3HDM MODEL

In the previous sections we have spelled out which conditions need to be satisfied to enforce axion-nucleon and axion-electron decoupling. Clearly, in a realistic scenario we expect that these conditions are realized only at some level of approximation, so that $C_{\mu, e, t}$ rather than vanish, will just be suppressed. We will now study more quantitatively the interrelation between nucleophobia and electrophobia, and the conditions to realise astrophobia with different levels of accuracy.

According to the results of the previous section, let us assume that the scalar potential contains the following terms:

$$H_3^* H_1 \phi^2 + H_2^* H_2 \phi^1.$$  

(19)

This corresponds to take $m = 2$ and $n = -1$ in Eq. (13) (first case). For the quarks we assume a $2 + 1$ structure with the PQ charges for the first generation equal to the ones of the second generation, as in model (1) in [1]. It is then sufficient to list the Yukawa operators involving just the second and third generations:

$$\bar{q}_2 u_2 H_1, \quad \bar{q}_3 u_3 H_2, \quad \bar{q}_2 u_2 H_1, \quad \bar{q}_3 u_3 H_2,$$

$$\bar{q}_2 d_2 H_2, \quad \bar{q}_3 d_3 H_1, \quad \bar{q}_3 d_3 H_2, \quad \bar{q}_3 d_3 H_1.$$  

(20)

We now assume that the leptons couple to a third Higgs doublet with the same charges for all generations:

$$\bar{\ell}_e e H_3.$$  

(21)

Equations (16) (17) imply $\chi_1 = \chi_3 - 2$ and $\chi_2 = \chi_3 + 1$. Neglecting flavor mixings, the diagonal axion couplings to quarks and leptons can be written as

$$C_{u,c} = \frac{2}{3} \frac{\chi_3}{3}, \quad C_t = \frac{1}{3} - \frac{\chi_3}{3},$$

$$C_{d,s} = \frac{1}{3} + \frac{\chi_3}{3}, \quad C_b = -\frac{2}{3} + \frac{\chi_3}{3},$$

$$C_{e, \mu, \tau} = \frac{\chi_3}{3}.$$  

(22)

Inserting the expressions for $C_{u,d}$ in the nucleophobic condition in Eq. (10) yields

$$\chi_3 = 1 - \frac{3}{2} f_{u_d} \approx -0.03,$$

(23)

which confirms that suppressed axion-electron couplings can indeed be compatible with nucleophobia ($C_e \approx -0.01$ to be compared, for example, with $C_e = 1/6$ for DFSZ models with $v_2 = v_1$).
The SN bound can be expressed as a constraint on the quantity $C_N = \sqrt{C_p^2 + C_n^2}$ [6,11] that, according to Eqs. (2)–(3), has a lowest value $C_N \approx 0.019$ which is determined by the correction $\delta_i$ in Eq. (2) (for comparisons in KSVZ axion models $C_N \approx 0.48$). To identify the parameter space regions corresponding to a sizeable suppression of the couplings, and to highlight the parametric correlation between nucleophobia and electrophobia, it is convenient to parametrize the VEVs as

$$v_1 = v c_1 c_2, \quad v_2 = v s_1 c_2, \quad v_3 = v s_2,$$  

where $s_i = \sin \beta_i$ and $c_i = \cos \beta_i$, and plot the value of $C_N$ and $C_e$ as a function of the angles $\beta_{1,2}$ rather than in terms of the VEVs ratios $\tan \beta_1 = v_2/v_1$ and $\tan \beta_2 = \sqrt{v_3^2/(v_2^2 + v_1^2)}$. Although the latter are the relevant physical parameters, this has the virtue of zooming in on the regions in which the values of the VEVs are not strongly hierarchical ($\tan \beta_{1,2} \sim O(1)$) and put in evidence the correlation between electrophobia and nucleophobia. In the parametrization of Eq. (24), the orthogonality condition in Eq. (15) reads $X_2 = (3c_1^2 - 1)c_2^2$. The requirement that the Yukawa couplings in the 3HDM remain perturbative restricts the allowed region in the $(\beta_1, \beta_2)$ plane. A conservative limit is obtained by imposing the tree-level unitarity bound on the $2 \to 2$ fermion scattering, $|\text{Re} a_{\mu,j=0}| < 1/2$, in the 3HDM theory involving Yukawas at $\sqrt{s} \gg M_{H_{1,2,3}}$. Ignoring running effects, which would put the constraint on $1/2$ in the parameter space regions corresponding to a sizeable suppression of the couplings.

We recall here for completeness that since nucleophobia requires generation-dependent PQ charges, axion couplings to quarks are in general flavor-violating. Thus, as discussed in [1], limits on FCNCs such as $K \to \pi\pi$ can also be used to constrain nucleophobic models. However, since these FCNC effects depend on the off-diagonal elements of the left and right unitary matrices that diagonalize the Yukawa couplings, which are not known, one can always assume a small mixing limit in which axion-related FCNC constraints are evaded.

A final remark about the axion coupling to photons is in order. The coupling is defined by the interaction term $y_{1,2} = 3\sqrt{2} \times 16\pi^3/3$ (from $Q_L \bar{u}_R \to Q_L \bar{u}_R$, with the initial and final states prepared into an $SU(3)_c$ singlet) and $y_{\tau}^{3\text{HDM}} < \sqrt{4/2\tau}$ (from $Q_L Q_L \to u_{\mu} \bar{u}_{\mu}$, with the initial state prepared into an $SU(2)_L$ singlet). The label 3HDM reminds us that these are not the Yukawa couplings of the SM. The latter are related to the former via: $y_1 = y_{1\text{HDM}} s_1 c_2$, $y_2 = y_{2\text{HDM}} c_1 c_2$, $y_3 = y_{3\text{HDM}} s_3$. The unitarity bounds on $y_{1,2,\tau}^{3\text{HDM}}$ can be now translated into a perturbativity bound in the $(\beta_1, \beta_2)$ plane, and this results in the hatched region in Fig. 1. Contour lines for different values of $C_N$ and $|C_e|$ are also plotted in Fig. 1, and show how electrophobia and nucleophobia occur in overlapping regions, so that a single choice of the values of the relevant parameters simultaneously realizes both properties. The two values $C_N = 0.05, 0.025$ labelling the black isolines in Fig. 1 correspond to the central value of light-quark mass ratio $m_u/m_d = 0.48 \pm 3$ [9]. Propagating the error one would obtain $C_N = 0.050 \pm 0.020$ and $C_N = 0.025 \pm 0.013$. We see from the figure that while for small values of $\beta_2$ the region with suppressed couplings is rather narrow, and astrophobia requires some tuning of the ratio $v_2/v_1$ to values sufficiently close to $\sqrt{2} (\beta_1 \approx 0.95)$, at larger values $\beta_2 \sim O(1)$ the region opens up and less tuning is required to simultaneously decouple the axion from nucleons and electrons.

We see from the figure that while for small values of $\beta_2$ the region with suppressed couplings is rather narrow, and astrophobia requires some tuning of the ratio $v_2/v_1$ to values sufficiently close to $\sqrt{2} (\beta_1 \approx 0.95)$, at larger values $\beta_2 \sim O(1)$ the region opens up and less tuning is required to simultaneously decouple the axion from nucleons and electrons.

5In multi-Higgs doublet models with no natural flavor conservation [14] scalars exchange can represent another source of FCNC. These effects can be easily avoided by assuming the decoupling limit [15] which yields in the low-energy spectrum a single neutral Higgs whose properties are indistinguishable from that of the SM-Higgs boson. This also ensures consistency of the model with LHC measurements of Higgs properties.

FIG. 1. Contour lines for $C_N = \sqrt{C_p^2 + C_n^2}$ (black) and $|C_e|$ (yellow) in the $(\beta_1, \beta_2)$ plane for the astrophobic model. For reference, contour lines corresponding to the values of $C_N$ and $C_e$ for the KSVZ and the DFSZ axion models are also plotted: $0.48 \to C_N^{\text{KSVZ}}$ (dotted line), $0.24, 0.66 \to |C_e|^{\text{DFSZ}}$ (grey region), $1/6 \to |C_e|^{\text{DFSZ}}$ (dotted yellow line) [6]. The light (dark) blue shaded area depicts the region where the axion mass bound from SN1987A is relaxed with respect to the KSVZ case by a factor of 10 (20).
where $C_p = E/N - 1.92(4)$, with $E$ denoting the coefficient of the electromagnetic anomaly. In astrophic models there is no particular reason for which $C_p$ should be suppressed. In the present model the contributions of the quarks and leptons are respectively $E_Q/N = 8/3 - 2X_3$ and $E_L/N = 2X_3$, so that their sum is $E/N = 8/3$, a value which is often encountered also in other axion models [16,17]. This suggests a neat experimental scenario in which the astrophobic properties of an axion might unambiguously emerge. Let us imagine that IAXO will discover the axion in a mass region $ma \gtrsim 0.20$ eV excluded, for standard axions, by the SN bound. This would be a compelling reason to believe that such an axion is nucleophobic. However, as we have argued, nucleophobia requires Higgs doublets charged under the PQ symmetry, and this in turn implies that the leptons are in general also PQ charged. Then, if the same mass value is also in tension with the white-dwarf and red-giants constraints, a sufficient suppression of the axion-electron coupling will be required as well, providing evidences for scenarios as the one we have discussed here. In this picture a final confirmation could eventually come from dedicated high-sensitivity searches for axion-related FCNC processes, since some amount of flavor violation in axion interactions with the quarks is a necessary ingredient for nucleophobia.

V. CONCLUSIONS

Astrophic axion models, wherein the axion couplings to nucleons and electrons can be simultaneously suppressed well below the values suggested by well-known benchmark models, can be elegantly implemented in a variant of the DFSZ model in which the PQ charges of the quarks are generation-dependent, and the scalar sector contains three Higgs doublets, one of which couples to the leptons and the other two to the quarks. Consistent astrophic axion models were first constructed in Ref. [1], however, in the original scenario axion decoupling from the electrons was achieved by means of a tuned cancellation between two different contributions to the axion-electron coupling, one proportional to the electron PQ charge, and the other generated by the mixing of the electron with the leptons of the heavier generations. The virtue of the 3HDM construction presented here is that it avoids the need for this cancellation, and enforces a strong correlation between the suppressions of the axion couplings to nucleons and electrons in such a way that nucleophobia and electrophobia are simultaneously realized in the same region of parameter space. This renders less contrived the possibility that axions might exhibit astrophic properties.

ACKNOWLEDGMENTS

F.B. and E.N. are supported by the INFN “Iniziativa Specifica” Theoretical Astroparticle Physics (TAstP-LNF). The work of L. D. L. is supported by the Marie Sklodowska-Curie Individual Fellowship Grant No. AXIONRUSH (GA 840791) and the ERC Grant No. NEO-NAT. F. M. is supported by MINECO Grant No. FPA2016-76005-C2-1-P, by Maria de Maetzu program Grant No. MDM-2014-0367 of ICCUB and 2017 SGR 929. F. M. acknowledges the INFN Laboratori Nazionali di Frascati for hospitality and financial support. L. D. L. and E. N. acknowledge hospitality from the Aspen Center for Physics (ACP) where this work was concluded. A. C. P. is supported by National Science Foundation Grant No. PHY-1607611. The participation of L. D. L. and E. N. at the A. C. P. was supported by the Simons Foundation.

[1] L. Di Luzio, F. Mescia, E. Nardi, P. Panci, and R. Ziegler, Phys. Rev. Lett. 120, 261803 (2018).
[2] M. Dine, W. Fischler, and M. Srednicki, Phys. Lett. 104B, 199 (1981).
[3] A. R. Zhitnitsky, Yad. Fiz. 31, 497 (1980) [Sov. J. Nucl. Phys. 31, 260 (1980)].
[4] J. E. Kim, Phys. Rev. Lett. 43, 103 (1979).
[5] M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, Nucl. Phys. B166, 493 (1980).
[6] M. Tanabashi et al. (Particle Data Group), Phys. Rev. D 98, 030001 (2018).
[7] J. H. Chang, R. Essig, and S. D. McDermott, J. High Energy Phys. 09 (2018) 051.
[8] P. Carenza, T. Fischer, M. Giannotti, G. Guo, G. Martinez-Pinedo, and A. Mirizzi, J. Cosmol. Astropart. Phys. 10 (2019) 016.
[9] G. Grilli di Cortona, E. Hardy, J. P. Vega, and G. Villadoro, J. High Energy Phys. 01 (2016) 034.
[10] F. Bjorkeroth, L. Di Luzio, F. Mescia, and E. Nardi, J. High Energy Phys. 02 (2019) 133.
[11] M. Giannotti, I. G. Irastorza, J. Redondo, A. Ringwald, and K. Saikawa, J. Cosmol. Astropart. Phys. 10 (2017) 010.
[12] L. Di Luzio, J. F. Kamenik, and M. Nardecchia, Eur. Phys. J. C 77, 30 (2017).
[13] L. Di Luzio and M. Nardecchia, Eur. Phys. J. C 77, 536 (2017).
[14] S. L. Glashow and S. Weinberg, Phys. Rev. D 15, 1958 (1977).
[15] H. E. Haber and Y. Nir, Nucl. Phys. B335, 363 (1990).
[16] L. Di Luzio, F. Mescia, and E. Nardi, Phys. Rev. Lett. 118, 031801 (2017).
[17] L. Di Luzio, F. Mescia, and E. Nardi, Phys. Rev. D 96, 075003 (2017).