This is the accepted manuscript made available via CHORUS. The article has been published as:

**Dark forces coupled to nonconserved currents**

Jeff A. Dror, Robert Lasenby, and Maxim Pospelov

Phys. Rev. D **96**, 075036 — Published 26 October 2017

DOI: [10.1103/PhysRevD.96.075036](https://doi.org/10.1103/PhysRevD.96.075036)
Dark forces coupled to non-conserved currents

Jeff A. Dror
Department of Physics, LEPP, Cornell University, Ithaca, NY 14853

Robert Lasenby
Perimeter Institute for Theoretical Physics, 31 Caroline Street N, Waterloo, Ontario N2L 2Y5, Canada

Maxim Pospelov
Perimeter Institute for Theoretical Physics, 31 Caroline Street N, Waterloo, Ontario N2L 2Y5, Canada and Department of Physics and Astronomy, University of Victoria, Victoria, BC V8P 5C2, Canada

New light vectors with dimension-4 couplings to Standard Model states have \((\text{energy/vector mass})^2\) enhanced production rates unless the current they couple to is conserved. These processes allow us to derive new constraints on the couplings of such vectors, that are significantly stronger than the previous literature for a wide variety of models. Examples include vectors with axial couplings to quarks and vectors coupled to currents (such as baryon number) that are only broken by the chiral anomaly. Our new limits arise from a range of processes, including rare \(Z\) decays and flavor changing meson decays, and rule out a number of phenomenologically-motivated proposals.

I. INTRODUCTION

New states beyond the Standard Model (SM) may have gone undetected either because they are too heavy to be produced in large numbers at collider experiments, such as those proposed by most solutions to the hierarchy problem, or because they couple very weakly to the SM states. In the latter case, such particles may be light enough to be produced in experiments other than the highest-energy colliders, and can have a diverse range of experimental signatures [1–3]. In this paper, we will focus on new light vector particles; these have been discussed extensively in the existing literature, for purposes including addressing experimental anomalies at low energies [4–9], explaining puzzles such as baryon stability [10], or acting as a mediator to a dark sector [11–13].

For a light vector with dimension-4 couplings to the Standard Model, unless the SM current that the vector couples to is conserved, there are processes with \((\text{energy/vector mass})^2\) rates involving the longitudinal mode of the new vector. In many such models, these energy-enhanced processes can be the dominant production mechanism in high-energy experiments, and can place strong constraints on the vector’s coupling. A number of works [5, 14–16] have used enhanced longitudinal production to place constraints on vectors with axial couplings to SM fermions. We extend these in a variety of ways. For vectors with axial, generation-non-universal, or \(SU(2)_L\)-violating couplings to SM fermions, we identify processes which yield stronger constraints than those in previous works. Most significantly, flavor-changing neutral current processes involving the new vector can be enhanced by \((\text{weak scale/vector mass})^2\) compared to competing processes. The resulting constraints are often the most powerful available, for vectors below the \(B\) mass.

We also point out that if the new vector couples to a current that is conserved at tree level, but broken by the chiral anomaly (for example the SM baryon number current), this anomalous non-conservation still gives rise to energy-enhanced longitudinal mode emission. These loop-level, but \((\text{energy/vector mass})^2\) enhanced, processes can place significantly stronger constraints on light vectors than existing ‘tree-level’ constraints. In the absence of fine-tuning, they can only be avoided at the expense of introducing extra sources of electroweak symmetry breaking in the UV theory, such as new sets of SM-chiral fermions. Such options generally run into strong experimental constraints. Conversely, cancelling the anomalies with new heavy fermions, that obtain their masses from a SM-singlet vacuum expectation value (VEV), always results in enhanced longitudinal emission. These points are also discussed, more concisely, in an accompanying letter [17].

Turning to the structure of the paper, in section II we discuss current non-conservation through the chiral anomaly, while in section III we treat more general non-conserved currents. Section IV describes how our new bounds can constrain several models of phenomenological interest. In particular, we show that several proposals to explain the \(^8\)Be anomaly [18] are effectively ruled out. Section V discusses some of the new experimental searches motivated by our analyses.

II. ANOMALOUS VECTORS

Even if the current that a vector couples to is conserved at tree level, it may be broken by chiral anomalies.
(examples in the SM are the baryon and lepton number currents). As a result, a vector $X$ coupled to such a current results in a non-renormalisable SM + $X$ effective field theory (EFT) — either the $U(1)_X$ symmetry is broken, or there is additional electroweak symmetry breaking (EWSB). In the former case, the UV cutoff scale of the theory is $\sim m_X/g_X$ (or below). This non-renormalisability is reflected in the divergence at high energies of some amplitudes in the EFT, involving fermion triangle diagrams (and box diagrams) with longitudinal $X$ modes. In this section, we review how to compute these amplitudes, and show how the anomaly results in enhanced longitudinal mode ($X_L$) emission in various circumstances.

This picture can appear puzzling from the perspective of a UV theory, in which the anomalies can be cancelled by new heavy fermions. The resolution is that, in addition to the fermion-mass-independent ‘anomalous’ 19 part of loop amplitudes, there is a piece that depends on the mass of the fermions in the loop. In the limit where the new anomaly-cancelling fermions have large masses compared to the external momenta around the loop, their contribution will correspond to that expected in the anomalous low-energy EFT, as we show explicitly below. Conversely, in the limit where their masses are small compared to the typical energies involved in the process, the mass-dependent piece vanishes, resulting in the usual anomaly-free result.

These considerations have been discussed in a number of previous papers, including [20–23], as well as the original papers of D’Hoker and Farhi [24, 25]. Our contribution in this section is essentially to present clearly the relation between the UV physics and the low-energy theory, and to point out that a $U(1)_X$-breaking theory results in energy-enhanced longitudinal production of a light $X$ via anomaly-couplings.

### A. Vectorial couplings

Since fermion masses introduce non-conservation of axial currents, a vector coupled to a tree-level-conserved current must have vectorial couplings to the SM fermions (and since the SM Yukawa couplings are non-diagonal, these couplings must be generation-universal). As the photon and gluon also have vectorial couplings, there is no mixed anomaly between the new vector $X$ and QED or QCD. However, the chiral anomaly will generically lead to non-conservation of the $U(1)_X$ current,

$$\partial^\mu J^\mu_X = \frac{A_{XBB}}{16\pi^2} \left( g^2 B_{\mu\nu} B^{\mu\nu} - g^2 W^a_{\mu\nu} (\tilde{W}^a)^{\mu\nu} \right), \tag{1}$$

where $A_{XBB} \equiv \text{Tr} [Q_X Y^2]$, and $A_{XWW} \equiv \text{Tr} [Q_X T^a T^a] = -A_{XBB}$ (since the current is vectorial), with the traces taken over the SM fermions and where $Q_X$ are their $U(1)_X$ charges. $V^{\mu\nu}$ are the field strength tensors, and $\tilde{V}^{\mu\nu} \equiv \frac{1}{2} \epsilon^{\mu\nu\sigma\rho} V_{\sigma\rho}$. If $A_{XBB}$ is non-zero, then the SM + $X$ EFT is not renormalisable, and some amplitudes will diverge at high energies.

In general, the effective theory may break the electroweak and $U(1)_X$ symmetries and hence we can include dimension-4 Wess-Zumino (WZ) terms,

$$\mathcal{L} \supset C_B g_X g^2 \epsilon^{\mu\nu\sigma\rho} X_\mu B_\nu \partial_\sigma B_\rho + C_W g_X g^2 \epsilon^{\mu\nu\sigma\rho} X_\mu (W^a_\nu \partial_\sigma W_\rho^a + \frac{1}{3} g^{abc} W^a_\nu W^b_\rho W^c_\sigma). \tag{2}$$

There are multiple ways of evaluating anomalous amplitudes within the effective theory, corresponding to different regularization schemes. The combination of a given regularization scheme, and particular values for the WZ coefficients, fixes the behaviour of the theory. To avoid breaking the electromagnetic gauge symmetry, we must have $C_B = -C_W$. As we will see below, these WZ terms can arise from integrating out heavy anomaly-cancelling fermions [24, 25]. If the WZ terms are not to introduce additional electroweak symmetry breaking — for example, if the heavy fermions get their mass from a SM-singlet vacuum expectation value — then their coefficient must be such as to cancel out the contribution of the $XBB$ and $XWW$ anomalies to the $W$ and $Z$ masses. Otherwise — for example, if the anomalies are cancelled by heavy SM-chiral fermions, which have been integrated out — the WZ coefficient may take other values.

### B. UV anomaly cancellation

If the anomalies are cancelled by heavy fermions, then for $U(1)_X$ to be preserved, the masses of these fermions must be EW-breaking. In the simplest case, they could obtain their masses through large Yukawa couplings with the SM Higgs. As reviewed in Appendix [A] this possibility is strongly constrained by electroweak precision tests and collider constraints; assuming that the LHC run-II sees no deviations from the SM, it will be fairly robustly ruled out. An extended EWSB sector would alter the details, but is generically also be subject to strong experimental constraints.

In UV completions where the heavy fermions have both SM-breaking and $U(1)_X$-breaking masses, the WZ coefficient in the low-energy theory depends on the relative size of these contributions. If the SM-breaking contributions are small compared to the total masses, then the WZ coefficient will be approximately that expected from a SM-preserving theory, up to $(m_{EW}/m_f)^2$ corrections. Conversely, if the dominant contribution to the masses is

---

1 EM gauge invariance also forbids a $eBW\partial W$ term. However, such a term will appear in the low-energy theory obtained by integrating out e.g. the top quark, in which the fermion content gives a chiral anomaly between the SM gauge bosons.
from an EW-breaking VEV, then the situation will be approximately that of the previous paragraph, and similar experimental constraints will apply.

A caveat to bear in mind is that such constraints rely on the existence of new, SM-chiral fermions, which have effects (such as electroweak precision observables) unsuppressed by the small coupling $g_X$. Within the low-energy theory, the effects of the SM-breaking WZ terms are all suppressed by $g_X$, and if this is small enough, may not be problematic. Consequently, it is possible that there exist more exotic $U(1)_X$-preserving UV completions, without anomaly-cancelling fermions, that are experimentally viable.

For the rest of this paper, we will focus on UV completions which result in a SM-preserving effective theory. These are easily realised; it is always possible to to introduce a new set of fermions with vectorial couplings to the SM gauge bosons, but chiral couplings to $X$, along with a $U(1)_X$-breaking VEV, to cancel the anomalies 26. As noted above, the lack of new EWSB fixes the coefficient of the WZ terms in the low-energy theory. While the value of the coefficient depends on the regularisation scheme chosen, the physical results are of course scheme-independent (see Appendix 13). Since the EFT breaks $U(1)_X$, these results include energy-enhanced emission of the longitudinal $X$ component, as we show below.

For UV completions with heavy anomaly-cancelling fermions, a slight complication is that the new ‘UV’ degrees of freedom do not necessarily have to be heavier than all of the SM states. For example, in the case of a vector coupled to the SM baryon number current, if we assume that anomalies are cancelled by SM-vector-like fermions, then collider constraints require that they have masses $\gtrsim 90$ GeV 27. If they are only slightly heavier than this bound, then for external momenta around the scale of the SM EW boson masses, the mass of the fermions in the loop will have an effect on anomalous $X_L$ amplitudes. In the following, we will assume that new states contributing to the anomaly are heavy enough that such momentum dependence can ignored, except where otherwise stated. This assumption will not be consistent for small enough $m_X/g_X$, since within the SM $+ X$ EFT, the growth of amplitudes with energy (as derived below) requires that there are new states at a scale $\lesssim 4\pi m_X/g_X$ (38). However, such large $g_X$ will generally be constrained more directly.

C. Triangle diagram amplitudes

To illustrate how the results outlined above arise, we will compute anomalous triangle amplitudes within the low-energy theory, and then show how this relates to the calculation in a UV-complete theory. Using the regularisation scheme from Appendix 13 that is symmetric between external legs, the longitudinal $XBB$ triangle amplitude is, summing over the SM fermions in the loop,

$$-(p+q)_\mu M_X^{\mu \rho} = \frac{A_{XBB}}{12\pi^2} g_X g^2 \epsilon^{\rho \lambda \sigma} p_\lambda q_\sigma,$$

As reviewed in Appendix 13 since the SM fermions have vectorial couplings to $X$, this amplitude does not depend on the masses of the SM fermions. This means that the momentum dependence of the longitudinal $X$ amplitude has the simple $\epsilon^{\rho \lambda \sigma} p_\lambda q_\sigma$ form, rather than involving extra terms depending on the external momenta compared to the mass of the fermions in the loop.

The regularisation scheme being symmetric between external legs means that we also have longitudinal $B$ amplitudes, $p_\nu M_B^{\nu \rho} = \frac{A_{XBB}}{12\pi^2} g_X g^2 \epsilon^{\nu \mu \lambda \sigma} q_\lambda p_\sigma$ etc (ignoring the SM fermion masses). To get rid of these, and restore the SM gauge symmetry within the SM $+ X$ EFT, we need an explicit Wess-Zumino term,

$$\mathcal{L} \supset \frac{A_{XBB}}{12\pi^2} g_X g^2 \epsilon^{\mu \nu \rho \sigma} X_B \partial_\nu B_\rho,$$

that gives a contribution to the amplitude of

$$-(p+q)_\mu M_W^{\mu \rho} = \frac{A_{XBB}}{6\pi^2} g_X g^2 \epsilon^{\rho \lambda \sigma} p_\lambda q_\sigma,$$

$$p_\nu M_W^{\nu \rho} = -\frac{A_{XBB}}{12\pi^2} g_X g^2 \epsilon^{\nu \mu \lambda \sigma} q_\lambda p_\sigma,$$

Adding together the contributions from the WZ term and from the SM fermion triangle diagrams, we obtain a total amplitude $M = M_X^{\mu \rho} + M_W^{\mu \rho}$ of

$$-(p+q)_\mu M^{\mu \rho} = \frac{A_{XBB}}{4\pi^2} g_X g^2 \epsilon^{\rho \lambda \sigma} p_\lambda q_\sigma,$$

with the SM gauge symmetry now preserved.

The motivation for adopting a symmetric regularisation scheme is that it makes clear how this amplitude, calculated within the SM $+ X$ EFT, relates to the calculation within a UV theory. The simplest UV completion, as discussed above, cancels the anomalies by introducing extra fermions which couple vectorially to the SM gauge bosons, but axially to $X$. These obtain heavy masses from a $U(1)_X$-breaking, but SM-singlet, VEV. In this setup, the ‘anomalous’ contributions to the $XBB$
amplitude cancel between the new fermions and the SM fermions (in any regularisation scheme), leaving only the fermion-mass-dependent pieces. Since the SM fermions have vectorial couplings to $X$, the mass dependence is only on the new fermions, which have axial $X$ couplings. The total longitudinal amplitude is

$$-(p + q)_\mu M^{\mu \nu} = \frac{1}{2\pi^2} \varepsilon^{\nu \rho \lambda \sigma} p_\lambda q_\sigma g X g g^2 \times$$

$$\sum_f 2m_f^2 I_{00}(m_f, p, q) X_{A,f} Y_f^2,$$

where $f$ runs over the new fermions, $Y_f$ is the hypercharge of $f$, $X_{A,f}$ is the axial coupling of $f$ to $X$, and the mass-dependent ‘scalar integral’ [19] term is

$$I_{00}(m_f, p, q) \equiv \int_0^1 dx \int_0^{1-x} dy \frac{1}{D(x, y, p, q)},$$

$$D \equiv y(1 - y)p^2 + x(1 - x)q^2 + 2xy p \cdot q - m_f^2.$$

For fermion masses well below the external momenta, i.e., $m_f^2 \ll p^2, q^2, p \cdot q$, the mass-dependent term vanishes, $2m_f^2 I_{00} \simeq 0$. This indicates that if there is no mass gap between the SM and heavy fermions, the longitudinal $X$ amplitude cancels between the two sectors. At the other extreme, if $m_f^2 \gg p^2, q^2, p \cdot q$ the mass dependent term approaches a constant, independent of heavy fermion mass, $2m_f^2 I_{00} \simeq -1$. The key here is that anomaly cancellation in the UV requires that

$$A_{XBB} = -2 \sum_f X_{A,f} Y_f^2$$

(11)

(the factor of 2 on the right-hand side comes from the fact that $f$ runs over the heavy Dirac fermions, each of which is made up of two Weyl fermions). Hence, if the masses of the heavy fermions are much greater than the external momenta then the total amplitude in equation [9] is

$$-(p + q)_\mu M^{\mu \nu} \simeq \frac{A_{XBB}}{4\pi^2} \varepsilon^{\nu \rho \lambda \sigma} p_\lambda q_\sigma g X g g^2,$$

giving the same result as the EFT calculation (equation [5]). We emphasize that this result is independent of the details of the UV theory, and only depends on it not introducing extra SM breaking. Equation [9] also illustrates how integrating out the heavy fermions gives the WZ term from equation [4] and how there will be $\sim p/m_f$ etc. corrections, corresponding to higher-dimensional operators within the effective theory.

The amplitudes for $XWW$ triangles will have similar behaviour, with $g'$ replaced by $g$. An additional feature is that, since $SU(2)_L$ is non-abelian, there are anomalous $XWW$ box diagrams, corresponding to the $XWWW$ part of the WZ term in equation [2]. These have an analogous story of fermion mass dependence in the UV theory.

Introducing gauge degrees of freedom for the longitudinal modes of the vector bosons, we could equivalently have calculated triangle diagrams between SM gauge bosons and the Goldstone mode for $X$. The latter has couplings $\propto m_f$ to the heavy fermions, so their contribution to triangle amplitudes becomes $\sim$ constant in the heavy $m_f$ limit (in exact analogy to fermions with large Yukawa couplings in the SM, as discussed in [23, 25]).

D. Axion-like behaviour

By the usual Goldstone boson equivalence arguments, the $1/m_X$-enhanced parts of amplitudes involving longitudinal $X$ are $\sim$ to those for the corresponding Goldstone (pseudo) scalar, $\phi$. Stated more precisely, we can take the limit $m_X \to 0, g_X \to 0$ with $f_X \equiv m_X/g_X$ held constant, decoupling the transverse $X$ modes. Then, making the substitution $g_X X_\mu \to \frac{f_X}{g_X} \partial_\mu \phi$ gives the same amplitudes in the $X$ and $\phi$ theories. For finite $m_X$, they will be equal up to $\mathcal{O}(m_X/E)$, where $E$ is some scale associated with the process.

In our case, the $X_L$ processes which survive in this limit all come from the anomalous couplings. In the $\phi$ theory, we can integrate by parts to write these couplings as

$$\frac{A_{XBB}}{16\pi^2} \frac{\phi}{f_X} (g'^2 W^\alpha \tilde{W}^\alpha - g'^2 B \tilde{B}) =$$

$$\frac{A_{XBB}}{16\pi^2} \frac{\phi}{f_X} \left( g^2 (W^+ \tilde{W}^- + W^- \tilde{W}^+) + gg' (\cot \theta_w - \tan \theta_w) Z \tilde{Z} + 2gg' Z \tilde{F} \right)$$

$$-ieg^2 \tilde{F}^{\mu \nu} (W_\mu^+ W_\nu^- - W_\mu^- W_\nu^+) \ldots \right),$$

(13)

where we have suppressed indices, and the dots correspond to further terms of the form $AW^+ W^-$ and $ZW^+ W^-$. Thus, energy-enhanced $X_L$ emission processes will have the same leading rate as the emission of an axion-like-particle (ALP) with these SM gauge boson couplings. This means that we can use the same processes that are used to search for light ALPs to look for $X$.

As mentioned above, within the simplest kinds of UV-complete models, the origin of equation (13) can be traced to explicit Yukawa interactions of anomaly-cancelling fermions with a set of $U(1)_X$ complex Higgs fields, for which $\phi$ is the Goldstone mode. Integrating out these fermions gives the couplings in equation (13), in exact analogy with the axion literature.

---

2 Processes involving large energies in a particular frame may still involve small invariant energy scales for the anomalous couplings. For example, if $X$ has an anomalous coupling to photons, then the production of an on-shell $X$ from two high-energy on-shell photons involves energies $\sim m_X$ in the rest frame of the $X$, and is in fact forbidden by the Landau-Yang theorem. In such cases, the $X$ amplitude can be $\mathcal{O}(1)$ different from Goldstone one.

3 the $WWWW$ terms from $W_{\mu \nu} (W^\alpha \mu \nu)^2$ cancel, reflecting the lack of pentagon anomalies for an abelian vector [20].
Since there is no two-photon anomalous coupling (as \( X \) has vectorial couplings to the SM fermions), longitudinal emission processes involving sub-EW-scale momenta are suppressed. Consequently, the most important effects of the anomalous couplings arise either in high-energy collisions — for example, on-shell \( Z \) decays — or in virtual processes which can be dominated by large loop momenta, such as rare meson decays.

**E. \( Z \rightarrow \gamma X \)**

If \( m_X < m_Z \), then the \( \varphi Z \tilde{F} \) coupling in equation 13 gives rise to \( Z \rightarrow \gamma X \) decays, with width

\[
\Gamma_{Z \rightarrow \gamma X_L} \simeq \frac{|A_{XBB}|^2}{384\pi^5} g_X^2 v^2 \frac{m_X^3}{m_X^2},
\]

(14)

\[
= \frac{|A_{XBB}|^2}{384\pi^5} g_X^2 \frac{m_X^3}{f_X^2},
\]

corresponding to a branching ratio

\[
\frac{\Gamma_{Z \rightarrow \gamma X_L}}{\Gamma_Z} \simeq 10^{-7}|A_{XBB}|^2 \left( \frac{\text{TeV}}{f_X} \right)^2.
\]

(15)

The corresponding experimental signatures and constraints depend on how \( X \) decays. At small \( m_X \) and \( g_X \), the \( X \) decay length will be longer than the scale of the experiment, so will give a missing energy signature (at the small \( g_X \) we are interested in, \( X \) will generally not interact strongly enough to be detected by its scattering). If \( X \) escapes the detector or decays invisibly, then LEP searches [30, 31] for single photons at half the \( Z \) energy constrain this branching ratio to be \( \lesssim 10^{-6} \).

For visible decays, there are published branching ratio limits for \( X \rightarrow \text{jet} + \text{jet} \) and \( X \rightarrow l^+l^- \) of \( \lesssim 3 \times 10^{-3} \) and \( 5 \times 10^{-4} \) respectively, with some improvement in the high mass region [34]. While these limits (from LEP) are not particularly stringent, we expect that the LHC has the capacity to significantly improve them. As illustrated by searches such as [35, 36], it is likely that leptonic \( Z \rightarrow \gamma (X \rightarrow l^+l^-) \) decays could be probed down to \( \mathcal{O}(10^{-5}) \) branching ratios or better.

Enhanced \( X_L \) emission also occurs for processes involving off-shell EW gauge bosons, with higher-energy processes giving increasing enhancements up to the scale of new states. For example, the rate for high-energy annihilation of SM fermions, \( \psi \bar{\psi} \rightarrow \gamma^* \rightarrow ZX \), will be enhanced by \( (E/m_X)^3 \), where \( E \) is the center of mass energy of the colliding particles. In some models, off-shell anomalous production may be the most promising LHC search channel, e.g. when \( X \) decays invisibly or escapes from the detector.

**F. FCNCs**

The coupling of \( X \) to quarks and the anomalous \( XWW \) coupling both lead to flavor changing neutral current (FCNC) interactions. Since the most important effects of these are at meson energy scales, the simplest procedure is to integrate out EW-scale states to obtain an effective \( Xqq' \) vertex. The QCD coupling is small at scales \( \sim m_W \), so the calculation is under perturbative control (see e.g. [37]). The leading effective interaction introduced is

\[
\mathcal{L} \supset g_{Xd_i}X_i d_j \gamma \mu \mathcal{P}_L d_i + \text{h.c.} + \ldots,
\]

(16)

where we have taken a down-type FCNC for illustration, and have omitted other, higher-loop-order diagrams (as well as \( X \) emission from external quark legs). The solid \( XWW \) vertex indicates the sum of WZ terms and fermion triangles (within a UV theory, it would simply be the sum over triangles). If \( X \) is coupled to a fully-conserved current, then \( g_{Xd_i} = 0 \), and the effective interaction is higher-dimensional; if \( X \) is coupled to a tree-level conserved current (as we consider here), then only the anomalous \( XWW \) coupling contributes to \( g_{Xd_i} \).

This effective operator then gives rise to flavor-changing meson decays. For heavy-quark decays, such as \( b \rightarrow sX \), these rates depend on hadronic matrix elements, which can be derived from QCD light-cone sum rules [35, 39]. For kaon decays, we can use chiral perturbation theory to obtain the leading approximation to the decay rates. Since the renormalisation of the left-handed quark current is proportional to the quark masses, the RG evolution of the effective operator from scales \( \sim m_W \) to meson energy scales can sensibly be ignored, in a first approximation. This interaction leads to meson decay rates through \( X_L \) emission of

\[
\Gamma(B \rightarrow KX) \simeq \frac{m_X^3}{64\pi m_X^2} |g_{bsX}|^2 \left( 1 - \frac{m_B^2}{m_K^2} \right)^2 |f_K(m_X^2)|^2 \frac{2Q}{m_B},
\]

(17)

\[
\Gamma(B \rightarrow K^*X) \simeq m_B^3 \left( |g_{bsX}|^2 |f_{K^*}(m_X^2)|^2 \left( \frac{2Q}{m_B} \right)^3 \right),
\]

(18)

\[
\Gamma(K^+ \rightarrow \pi^0 X) \simeq \frac{m_{K^+}^3}{64\pi m_X^2} \left( 1 - \frac{m_{K^+}^2}{m_{K^+}} \right)^2 |g_{sdX}|^2 \frac{2Q}{m_{K^+}},
\]

(19)

\[
\Gamma(K_L \rightarrow \pi^0 X) \simeq \frac{m_{K_L}^3}{64\pi m_X^2} \left( 1 - \frac{m_{K_L}^2}{m_{K_L}} \right)^2 \text{Im}(g_{sdX})^2 \frac{2Q}{m_{K_L}},
\]

(20)

where \( Q \) is the momentum of the decay products in the center of mass frame, and \( f_K, f_{K^*} \) are the appropriate form factors coming from the hadronic matrix elements.
(we use the fits from \cite{38,39}). The deviation of the kaon decay rate from the leading chiral perturbation theory value given above is expected to be of order a few percent (see e.g. \cite{40}). Since we are quoting the leading $1/m_X^2$ rates, we would obtain same-order results by evaluating the form factors at zero momentum transfer, $f_K(m_X^2) \simeq f_K(0)$ etc.

In the calculation of $g_{d^3d^3X}$, while each individual diagram in \cite{16} is divergent, these divergences cancel in the sum over virtual up-type quarks. This occurs since the divergent terms are independent of the quark mass, so their sum cancels due to the unitarity of the CKM matrix. As a result, the integral is dominated by moments $\sim m_t$, and couplings suppressed by the cutoff scale will give sub-leading contributions (in the UV theory, the masses of the UV fermions in triangles will be much larger than the external momenta of these triangles). The coefficient of the effective vertex is

$$
g_{Xd^3d^3} = \frac{-3g^4A_{XBB}}{16\pi^2} - g_X \sum_{\alpha \in \{u,c,t\}} V_{\alpha d}V_{\alpha d}^* F \left( \frac{m_\alpha^2}{m_W^2} \right) + \ldots ,
$$

(21)

where

$$
F(x) \equiv \frac{x(1 + x(\log x - 1))}{(1 - x)^2} = x + O(x^2 \log x) .
$$

(22)

Due to the $m_q^2/m_W^2$ dependence for small quark mass, the sum over up-type quarks is dominated by the top quark, for both $bsX$ and $sdX$ vertices, despite the smallness of the $V_{ts}$ and $V_{td}$ CKM matrix elements. Since $m_t^2/m_W^2$ is small, the equivalent up-type FCNC vertices, such as $cuX$, are suppressed compared to down-type FCNCs.

Compared to these effective FCNC vertices, other effective flavor-changing operators are higher-dimensional, and so are suppressed by more powers of $g_X/m_X$ and/or $1/m_X^2$. Thus, despite equation (21) representing a 2-loop contribution (within the UV theory), it is able to dominate over 1-loop $d^3d^3X$ processes. For example, in the $B \to KX$ decay we have,

$$
\mathcal{M}^2 - \text{loop} / \mathcal{M}^2 - \text{loop} \propto g^2/(16\pi^2) \times (m_t/m_X)^2 ,
$$

which, for $m_X$ light enough to be emitted in the decay, is $\gg 1$. \footnote{The $\propto m_X^2$ (rather than $\propto m_X$) relative suppression of 1-loop emission comes from angular momentum conservation in the pseudoscalar $\to$ pseudoscalar + vector decay; for $B \to K^*X$ decays, we would have $\mathcal{M}^2 - \text{loop} / \mathcal{M}^2 - \text{loop} \propto m_t^2/(m_X m_b)$ instead.} Competing SM FCNC processes are also suppressed; for example, the $bs\gamma$ vertex is of the form $\propto m_b^2 F_{\mu\nu} b_L \gamma^{\mu\nu} s_L + \ldots$ (on-shell), since the photon couples to a conserved current, while 4-fermion vertices are suppressed by at least $G_F$.

If $m_X$ is light enough, then FCNC meson decays via an on-shell longitudinal $X$ become possible, and are enhanced by (energy/$m_X$)$^2$, in addition to being lower-dimensional than other effective flavor-changing processes. Most directly, the $bsX$ and $sdX$ vertices result in $B \to K^{(*)}X$ and $K \to \pi X$ decays, giving new flavor-changing meson decays that can place strong constraints on the coupling of $X$. This is in exact analogy to the FCNC processes discussed in \cite{43}, for axion-like particles with a coupling to $W^aW^a$. In contrast, processes involving two or more $d_dX$ vertices, such as the $X$ contribution to meson oscillations, are suppressed by $1/f_X^2$, but compete with SM processes suppressed by $1/m_X^2$. Consequently, it is difficult for such processes to probe $f_X$ above the EW scale, unless $m_X$ is accidentally very close to the meson mass, resulting in resonant enhancement from meson-$X$ mixing.

The selection rules for decays via longitudinal vector emission are different to those for transverse emission. In the latter case, angular momentum conservation suppresses (pseudo)scalar $\to$ (pseudo)scalar + vector decays, since these demand that the vector’s spin is perpendicular to its momentum. This suppresses the rate of such decays via a vector that couples to a conserved current; for example, there are no $B^+ \to K^+\gamma$ decays, while the rate for $B^+ \to K^+A'$, where $A'$ is a kinetically-mixed dark photon, is proportional to $m_{A'}^2$. However, by Goldstone boson equivalence, meson decays via a light longitudinal $X$ have the same rates as the corresponding ALP decays, so decays such as $B^+ \to K^+X$ are unsuppressed.

1. Experimental constraints

Here, we summarise the experimental searches we will use to constrain FCNC meson decays via $X$ — for easy reference, these are tabulated in Table I.

If $X$ is sufficiently light and weakly coupled that it decays outside the detector, then $B \to K\nu\bar{\nu}$ and $K \to \pi\nu\bar{\nu}$ searches constrain the $B \to KX$ and $K \to \pi X$ branching ratios. The $K \to \pi\nu\bar{\nu}$ channel is especially constraining, with existing experiments having measured a very small ($\sim 10^{-10}$) branching fraction consistent with the SM prediction \cite{53,54}, which the future NA62 experiment should be able to measure to $\sim 10\%$ relative error \cite{55}.

For prompt decays of $X$ into leptons, as can occur for heavier / more strongly coupled $X$, searches for $B \to K^{(*)}\ell^+\ell^-$ and $K \to \pi^{(*)}\ell^+\ell^-$ decays place strong constraints. The LHCb search for $B^+ \to K^+\ell^+\ell^-$ measures the branching ratio to be $(4.36 \pm 0.15 \pm 0.18) \times 10^{-7}$ \cite{46}. For kaons, the $K^+_L \to \pi^0 e^+e^-$ decay is very well-constrained, with a branching ratio bound of $\lesssim 3 \times 10^{-10}$ \cite{50}. However, because of the large hadronic branching ratios for $K^0 \to \pi^0\pi^0$ and $K^0_L \to \pi^0\pi^0$, the Dalitz decay $\pi^0 \to e^+e^-$ gives a background that makes $K^0 \to \pi^0 e^+e^-$ measurements difficult at $m_{\ell\ell} \lesssim m_{\pi^0}$ \cite{50} (the same applies to $K^\pm \to \pi^\pm e^+e^-$ versus $K^\pm \to \pi^\pm \pi^0$ \cite{56}). Thus, for $m_X \lesssim m_{\pi^0}$, the best constraints come from $B \to K^{(*)}e^+e^-$ decays, where the competing $B \to K\pi^0$ decays are also suppressed. For example, the $B \to K^*e^+e^-$ branching ratio is measured to be $\simeq 10^{-6}$.
for $m_{ee} \lesssim 300$ MeV [49].

If $X$ dominantly decays into hadrons, then it may be possible to perform bump-hunt searches in the invariant mass distributions of $B \to K +$ hadronic decays. For example, the $B \to K \omega$ decay is detected as a peak in the $m_{3\pi}$ distribution of $B \to K \pi^+ \pi^- \pi^0$ decays, with branching ratio error $\sim 10^{-6}$ [57]; a similar search could be performed at other invariant masses.

In addition to the prompt and invisible decays discussed above, it is also possible to look for displaced $X$ decays. LHCb performed a search for long-lived scalar particles decaying into $\mu^+ \mu^-$ in $B \to KX$ decays [48]. These limits dominate in the displaced regime with branching ratio limits reaching $\sim 2 \times 10^{-10}$. For very displaced decays, the best constraints come from beam dump experiments. Here, the enhanced $K \to \pi X_L$ decay means that kaon decays, which are usually a sub-dominant production mechanism in proton beam dump experiments (for tree-level vector couplings), can be the dominant process through which $X$s are produced. This allows proton beam dump experiments such as CHARM [58] and (in the future) SHiP [59] to probe smaller couplings than indicated by a naive analysis [59] [60].

It should be noted that, unlike constraints involving visible $X$ decays, missing energy searches are effective down to arbitrarily small vector masses, and constrain correspondingly tiny $g_X$ for small $m_X$. However, for $X$ with couplings to first-generation fermions, the strong constraints coming from stellar energy loss bounds [61] [62], and from fifth force / equivalence principle tests at smaller $m_X$ [63], mean that it is generically only at extremely small $m_X$ that missing energy constraints become the dominant bound.

### Table I. Summary of the FCNC searches used to constrain models with a light vector $X$, in the text and in figures.

| Decay type | Measured branching ratio | Reference | Comments |
|------------|--------------------------|-----------|----------|
| $B \to K (X \to \ell^+ \ell^-)$ | $(4.7 \pm 0.6 \pm 0.2) \times 10^{-7}$ | BaBar, 2012 [44] | $m_{\ell\ell} > 0.1$ GeV$^2$ |
| $B \to K (X \to \text{inv})$ | $< 1.6 \times 10^{-5}$ | Belle, 2017 [45] |
| $B^+ \to K^{+}(X \to \mu^+ \mu^-)$ | $(4.36 \pm 0.15 \pm 0.18) \times 10^{-7}$ | LHCb, 2012 [46] | $m_{\mu\mu} > 0.05$ GeV$^2$ |
| $B^0 \to K(X \to 3\tau)$ | $< 2.3 \times 10^{-4}$ | Particle Data Group [47] |
| $B^0 \to K(X \to \mu^+ \mu^-)$ | $< 2 \times 10^{-10} - 10^{-7}$ | LHCb, 2016 [48] | displaced search |
| $B \to K^*(X \to \ell^+ \ell^-)$ | $(10.2^{+1.4}_{-1.3} \pm 0.5) \times 10^{-7}$ | BaBar, 2012 [43] | $m_{\ell\ell} > 0.1$ GeV$^2$ |
| $B \to K^*(X \to \text{inv})$ | $< 1.6 \times 10^{-5}$ | Belle, 2017 [45] |
| $B^+ \to K^{*+}(X \to e^+ e^-)$ | $(1.32^{+0.41}_{-0.36} \pm 0.09) \times 10^{-6}$ | BaBar, 2008 [49] | $m_{ee} < 0.1$ GeV |
| $B^0 \to K^{0}(X \to e^+ e^-)$ | $(0.73^{+0.22}_{-0.19} \pm 0.04) \times 10^{-6}$ | BaBar, 2008 [49] | $m_{ee} < 0.1$ GeV |
| $K_L \to \pi^0(X \to e^+ e^-)$ | $< 2.8 \times 10^{-10}$ | KTeV/E799, 2003 [50] |
| $K_L \to \pi^0(X \to \mu^+ \mu^-)$ | $< 3.8 \times 10^{-10}$ | KTeV, 2000 [51] |
| $K^\pm \to \pi^\mp(X \to \mu^+ \mu^-)$ | $< 1.1 \times 10^{-9}$ | NA48/2, 2011 [52] |
| $K^+ \to \pi^+(X \to \text{inv})$ | $(1.73^{+1.15}_{-1.05}) \times 10^{-10}$ | E949, 2008 [53] |

G. Baryon number coupled vector

To give an example of how these constraints relate to each other and to other bounds in the literature, for a specific model, we will consider a vector coupled to the SM baryon number current. This model has been investigated in many papers over the past decades, with motivations including acting as a stabilisation mechanism for baryon number [10], mediating a new force that avoids the strong constraints coming from leptonic, axial, or non-Minimal Flavor Violation couplings [54], as well as addressing experimental anomalies [9].

The SM baryon number current is conserved at tree level, but broken by hypercharge and $SU(2)_L$ anomalies, with $A_{XBB} = -A_{XWW} = n_g / 2$, where $n_g = 3$ is the number of SM generations. Consequently, subject to caveats regarding the UV completion (see below), there will be anomalous production processes of the kind we have considered above.

Assuming that $X$ does not decay to hidden sector states, then for $m_X \gtrsim m_{ee}$, the vector will dominantly decay to hadronic states. While a ‘pure’ baryon-number-coupled vector does not have tree-level couplings to leptons, a kinetic mixing with the photon is generated by RG evolution, so can only be set to zero at a single scale. For commonality with other literature, such as [64], we will assume a kinetic mixing $\epsilon = e g_X / (4\pi)^2$ for our plots. Consequently, for $2m_e < m_X \lesssim m_{ee}$, $X$ will decay to $e^+ e^-$ through the kinetic mixing.
processes. Here we are assuming a UV completion that preserves the EW symmetry in order to evade electroweak precision and direct collider constraints (see section [11B]). As this figure indicates, including the anomalous processes gives a significant improvement in the constraints across a wide mass range. To reiterate, these anomalous-coupling-based bounds arise whenever the mixed electroweak-$U(1)_X$ anomalies are cancelled by heavy fermions whose masses do not (dominantly) arise from EW-symmetry breaking. The limits displayed in Figure 1 correspond to the UV completion introducing no extra EW-symmetry breaking; if the heavy fermions receive some fraction of their mass from an EW-symmetry breaking contribution, then the limits will be reduced proportionately. Significantly weakening these bounds would require either more exotic UV completions which do not introduce new fermions, or for the new fermions to have dominantly EW-breaking masses, with the latter option running into strong observational constraints (see e.g. Appendix A).

For the couplings shown in Figure 1, the decay time of $X$ is $<< 1$ sec for $m_X > 2m_e$, so for $m_X \gtrsim 10$ MeV, $X$ will have a strongly Boltzmann-suppressed abundance in the early universe at temperatures $\lesssim 2$ MeV, when non-equilibrium processes are important. For $m_X < 2m_e$, the lifetime will be long enough for there to be constraints from the late-time decay of an $X$ abundance produced in the hot early universe, while for $2m_e < m_X \lesssim 10$ MeV, there will be constraints, at large enough $g_X$, from $X$ freeze-out occurring after SM neutrino freeze-out [65]. We leave the calculation of such constraints to future work. Supernova cooling bounds [61] will also constrain some range of parameters; however, the emission rates computed in the literature [60, 61] do not take into account either in-medium mixing effects [62] or anomalous production, so would need to be re-calculated.

The right-hand panel of Figure 1 shows the constraints that arise if $X$ has a significant branching ratio to invisible states. For example, one light Dirac fermion $\chi$ with $X$-charge of 1 and $2m_e < m_X$ will result in an invisible branching fraction of $\gtrsim 30\%$. The constraints from missing energy searches are strong, and limit the discovery prospects for light dark matter coupled through such a mediator at neutrino experiments [68, 69]. A small modification of this plot will also apply to a vector coupled to lepton number. This has the same anomalous couplings as a baryon number vector, but its branching ratio to neutrinos is larger than $30\%$ everywhere (and is $\sim 100\%$ for $m_X < 2m_e$). In addition to the anomalous bounds, a lepton number coupled vector will also be constrained by neutrino-electron scattering experiments (section 11F), which provide the dominant constraints for $m_X \gtrsim m_K - m_\pi$.

1. Comparison to literature

Some early papers on baryon number vectors, such as [10, 70], considered models in which the anomalies are cancelled by new SM-chiral fermions, and noted that these could be made compatible with the electroweak precision (EWP) constraints of the time. At least one later paper [71] considered $X$ production through its anomalous couplings to SM gauge bosons (presumably assuming a model where anomalies are cancelled at a higher scale), but performed the calculation incorrectly [6].

More recent papers, such as [49, 64], consider models in which the anomalies are cancelled at a higher scale in a SM-preserving way, but ignore the anomalous couplings in the low-energy theory. [27] considers such a model, and studies the direct production of the heavy fermions at colliders. The constraints derived are, at $m_X \lesssim m_K$, significantly weaker than those from the anomalous couplings present in such models. This is in analogy to searches for axion-like particles, in which it is often easier to detect effects involving the light degree of freedom than it is to probe the heavy states beyond the cutoff.

A number of works have looked at the consequences of ‘effective Chern-Simons’ interactions of a new vector with the SM, i.e. couplings of the form $e^{\omega \mu \nu} X_{\mu} Z_\nu \partial_\lambda A_\lambda$ etc to the SM EW gauge bosons, as reviewed in [59]. Some of these works, such as [72] and [59], seem to claim that unsuppressed SM-violating values for these couplings can occur in the SM + $X$ EFT, due to heavy anomaly-cancelling fermions which dominantly obtain masses from a SM-singlet VEV. As we have discussed, the values of such WZ terms are determined, up to $m_{EW}^2/m_Z^2$ corrections, by the couplings of $X$ to SM fermions, within that class of UV completions. Most works other than [59] also consider $m_X > m_Z$, giving rather different phenomenology. For low $m_X$, the FCNC constraints we have derived dominate those given in [59].

An interesting side point is that the kind of anomalous interactions we have considered for a new light vector may also have analogues within the SM. [73] propose that the presence of the $\omega$ meson in the low-energy theory of QCD, which behaves analogously to a baryon number coupled vector, should require a WZ term of the form

\[ \text{[6] In particular, [71] states that “The contribution of a fermion does not depend on its mass while the latter is much smaller than $m_Z$, and its contribution is suppressed with factor } \lesssim 0.1(m_Z/m_\tau)^2 \text{ for } m_\tau \gtrsim m_Z (\text{so we do not take the top quark into account}). ” As per section 11C, since the $X$ couples vectorially to SM fermions, the amplitude for longitudinal $X$ emission does not depend on the mass of the SM fermions in the triangle diagrams. Ref. [77]’s calculations would, for example, imply an anomalous amplitude for longitudinal $B-L$ vector emission, which cannot be present.} \]
\[ \epsilon^\mu_\nu \omega_\mu Z_\nu F^\rho_\sigma, \] which in the low-energy theory would result in a \( \omega \gamma \bar{\nu} \nu \) coupling. Since the \( \omega \) meson is composite at scales comparable to its mass, high-energy \( E^2/m^2_\omega \) enhanced processes of the type we have considered will not apply.

### III. TREE-LEVEL BREAKING

In addition to being broken through chiral anomalies, as considered in the previous section, the SM current that a light vector is coupled to may also be broken at tree level, within the SM + X EFT. The possible sources of breaking via dimension-4 couplings are:

- **Axial couplings to SM fermions**: since fermion Dirac masses break axial symmetry, an axially-coupled vector \( X \) will have longitudinal \( X \) production processes enhanced by \( (m_f/m_X)^2 \), where \( m_f \) is the mass of the relevant SM fermion.\(^7\) Axial couplings also allow \( X \) to have ‘anomalous’ two-photon and two-gluon couplings — in fact, as we will see in section III.C they render such couplings unavoidable, due to the dependence of triangle amplitudes on SM fermion masses. Consequently, processes involving heavy SM fermions, or anomalous couplings, can provide strong constraints on axial couplings.\(^7\)

- **Generation-non-universal couplings**: due to the non-diagonal quark mass matrices, generation-non-universal couplings of a vector to quarks generically lead to tree-level FCNCs (see Appendix [C]), which are tightly constrained by experiment. If the vector is light, then these vertices can lead to flavor-changing quark decays via the emission of a real longitudinal \( X \), placing even stronger constraints on such operators. Even without tree-level FCNCs, generation-non-universal couplings in combination with SM flavor-changing vertices generally result in \( X \)-non-conserving processes (section III.B). Lepton flavor-changing vertices are also subject to strong experimental constraints, but can be more easily avoided. Given the strong experimental constraints on tree-level flavor-changing couplings, we will assume throughout this paper that these are highly suppressed.

- **Weak-isospin violation**: \( W^q_\mu q_d \) or \( W^\ell \bar{\nu} \) vertices may be \( U(1)_X \)-breaking, if \( X \) couples differently to left-handed fermions that are members of the same \( SU(2)_L \) doublet, and does not have the compensating coupling to the \( W \). \( X \) radiation from charged current processes is then enhanced.

- **EW couplings**: in addition to its couplings to fermions, \( X \) may have dimension-4 couplings to the SM EW gauge bosons, or to the SM Higgs. A simple example is mass mixing with the Z boson \([76, 77]\), which leads to a \( XWW \) coupling. For simplicity, we will only consider Wess-Zumino type couplings in this paper; as discussed above, these are determined by the SM fermion couplings of \( X \), if the UV completion does not introduce extra EWSB, and are suppressed by a loop factor.

In this section, we will identify various process which place strong constraints on couplings of these forms. To keep the discussion as general as possible, we parameterize the interaction between \( X \) and SM fermions as

\[ \mathcal{L} \supset g_X X_\mu \sum_i \bar{\psi}_i \gamma^\mu (c^Y_i + c^A_i \gamma_5) \psi_i. \] (24)

#### A. Radiation from axial current

The simplest such process is the radiation of a longitudinal \( X \) from a massive SM fermion via its axial coupling. This is enhanced by \( (m_f/m_X)^2 \), so it is advantageous to use heavy quarks or leptons in these searches.

As a relevant example, consider heavy quarkonium decays such as \( \Upsilon \rightarrow \gamma X \). Here, the coupling of \( X \) to the \( b \) quark leads to enhanced \( X \) emission. This process has been considered in \([14, 78]\) (and in \([79]\) for the case of a light pseudoscalar), and has a branching ratio of

\[ \text{Br}(\Upsilon_{1S} \rightarrow \gamma X) \simeq 4 \times 10^{-5} |c^b_1|^2 \left( \frac{\text{TeV}}{m_X} \right)^2. \] (25)

The \( \Upsilon \rightarrow \gamma + \text{invisible} \) \([80]\) and \( \Upsilon \rightarrow \gamma (X \rightarrow \mu^+ \mu^-) \) \([81]\) decays are both measured to have small \((<10^{-5})\) branching ratios, giving constraints on \( g_X c^A_1 \). Similar constraints can be formulated for the axial charm coupling through \( J/\psi \) decays (using, e.g., constraints from \([82]\) ).

An axial coupling to the top quark gives the largest \( m^2_\tau/m^2_X \) enhancement. However, the much smaller number of top quarks produced in experiments means that, at small \( m_X \), axial couplings to lighter quarks will give stronger constraints, while at larger \( m_X \), other high-energy production mechanisms will generally dominate.

As we discuss in the next section, the constraints on quark axial couplings from FCNC meson decays generally dominate those from other processes, such as \( \Upsilon \) decays, except in the case of purely right-handed down-type couplings.

---

\(^7\) The QCD chiral condensate also breaks chiral symmetry, so axial coupling processes involving hadrons such as nucleons are not suppressed by the small first-generation quark masses, but only by the nucleon mass etc.

\(^8\) If SM neutrinos are Majorana, then there will also be \( U(1)_X \) breaking effects from couplings to neutrinos, suppressed by the small neutrino masses; we will ignore these.
is broken at tree level, then the effective up-type quark loop. If constraints are (left to right) from φ approximate region in which anomaly-cancelling fermions would be light enough to have been detected [27]. The other gray

gograms involving the X d cels, there must be flavor-changing

Consequently, unless the sum of these divergences can-

W couplings to quarks [83], compared to one with a

analogy to meson decays via an axion-like particle with

from flavor-changing meson decays. This is in precise

means that even stronger constraints can be obtained

in section II F, the lack of additional loop suppression
to the currents broken by the chiral anomaly considered

in the EFT. Thus, unlike in section II F, we cannot as-

This prescription is complicated by the fact that the divergent parts of EFT amplitudes are not gauge-
independent; however, their $m_t^2/m_W^2$-enhanced parts

B. FCNCs

As reviewed in section II F flavor-changing transitions between down-type quarks can proceed via a W-boson / up-type-quark loop. If X couples to a quark current that is broken at tree level, then the effective $d_i d_j X$ vertex obtains a value $\sim g_X^2 g Y (\text{CKM elements})$. Compared to the currents broken by the chiral anomaly considered in section II F the lack of additional loop suppression means that even stronger constraints can be obtained from flavor-changing meson decays. This is in precise analogy to meson decays via an axion-like particle with couplings to quarks [82], compared to one with a WW coupling [83].

Within the SM + X EFT, flavor-changing penguin diagrams involving the X coupling to quarks are divergent. Consequently, unless the sum of these divergences cancels, there must be flavor-changing $d_i d_j X$ counterterms in the EFT. Thus, unlike in section II F we cannot assume that the UV theory contributes no unsuppressed $d_i d_j X$ FCNCs. However, we can estimate the contributions from 'simple' UV completions by evaluating the logarithmically divergent $d_i d_j X$ amplitudes within the EFT, and assuming that these are resolved by UV states, giving an overall $\log(M/m_{EW})$ amplitude (where M is the UV scale).

This prescription is complicated by the fact that the divergent parts of EFT amplitudes are not gauge-independent; however, their $m_t^2/m_W^2$-enhanced parts

$$g_X^4 \frac{d_i d_j}{g_2 g_Y} C_{d_i d_j} V_{td} V^*_d + \ldots,$$

$C_{d_i d_j} = \frac{1}{2} \frac{m_t^2}{m_W^2} (c_{d_i}^L + c_{d_j}^L - 2 c_{d_i}^R) \log \frac{M^2}{m_t^2} + \ldots$

(where we have neglected the down-type quark masses).

In the case of universal vectorial couplings, this vanishes, as expected from a coupling to a conserved current. Below, we will see how an effective coupling of this kind arises in a two Higgs doublet model (2HDM) UV completion, where the UV states are the heavy charged Higgses. It is of course possible that other UV completions may lead to further cancellations; unlike for the anomalous XWW coupling, evaluating FCNC amplitudes properly requires knowing the full theory.

Interestingly, the FCNC vertex can be enhanced by the top quark mass (as opposed to suppressed by a light quark mass) even in the case where X does not couple to the top quark at all. For example, left-handed couplings to down-type quarks can also give a $m_t^2/m_W^2$ enhanced vertex, due to the self-energy diagrams.

9 This is only within the family of $R_t$ gauges; the situation is more complicated in unitary gauge. In fact even within the SM, off-shell $d_i d_j Z$ amplitudes are divergent in unitary gauge, with divergences only cancelling when combined with $W^+ W^-$ box diagrams, or when the Z is put on-shell [84].
As discussed in Appendix C, generation-non-universal couplings to quarks can lead to tree-level flavor-changing vertices. In particular, if $X$ couples to both up- and down-type left-handed quarks, then either these couplings are related in a specific weak-isospin-violating way, or there are tree-level FCNCs. From equation (26), we see that unless the left-handed down-type couplings are equal, then there will be some down-type FCNCs which are unsuppressed by small quark masses. Thus, the safest way to have generation-non-universal couplings to quarks, without introducing dangerous FCNCs, is generally to have dominantly right-handed couplings.

Note that we assumed above that $X$ does not mix with the $Z$. If there is such a mixing, the resulting $XWW$ coupling gives an additional contribution to FCNCs — in Feynman gauge, the coupling of $X$ to the charged Goldstones will give a $m_b^2/m_W^2$ enhanced, logarithmically-divergent, contribution [11].

As discussed in section [11F.1] experimental searches for FCNC meson decays via these effective operators will depend on how $X$ decays. Compared to the effective operators arising from anomalous $XWW$ couplings, the lack of an extra loop suppression factor in equation (26) results in larger branching ratios for a given coupling, so gives stronger constraints on $g_X$.

1. UV completion

It can be enlightening to see how the effective vertex arises within a particular UV completion. Here, we calculate the $d_d_j X$ vertex within the 2HDM $U(1)_R$ UV completion presented in [77]. In their UV theory, $X$ has purely right-handed couplings to SM fermions, while there are Yukawa terms,

$$\mathcal{L} \supset Y_u H_u Q_u R + Y_d H_d Q_d R + Y_t H_d L_t R + \text{h.c.}, \quad (27)$$

along with new anomaly-canceling fermions which get masses from a pair of SM-singlet, but $X$-breaking, VEVs. Since $H_u$ and $H_d$ have $X$ charges, the SM fermions can gain comparable vector and axial couplings to $X$. As illustrated in Figure 2, calculating the $d_d_j X$ vertex in this model involves adding the charged Higgs exchange diagrams, which cancel the logarithmic divergences in the $W$ exchange diagrams. Keeping only the $m_t^2/m_W^2$ enhanced terms, and choosing the 2HDM parameters so that $X$ does not mix with $Z$ (i.e. $q_{H_u} s_{H_d}^2 = q_{H_d} c_{H_u}^2$, in the notation of [77]), we obtain the effective vertex, $g_{X d_d j}^{2HDM} = \frac{1}{16\pi^2} g_X v \frac{Y_{t \beta}}{c_{\beta}} \log \frac{m_{t \beta}^2}{m_H^2}.$ \quad (28)

Since the $u_R$ coupling of $X$ in this model is $C_u^R = \frac{1}{2} g_X q_{u_R} = \frac{1}{2} g_X (q_{H_u} + q_{H_d}) c_{\beta}^2 s_{\beta}^2$, this gives,

$$g_{X d_d j}^{2HDM} = -\frac{1}{16\pi^2} c_{\beta}^2 c_{H_d}^2 v g_X \frac{m_t^2}{m_W^2} \log \frac{m_{H_d}^2}{m_t^2}. \quad (29)$$

As expected, this matches the form of the log-divergent terms in the EFT calculation given by equation (26), with the UV scale set by the mass of the charged Higgs states. In a realistic 2HDM model, this scale may not be so far above the top mass (see e.g. [85]), and non-log-enhanced terms may be numerically important. However, in the absence of fine-tuned cancellations, the log-enhanced term should be a good parametric estimate of the effective vertex, and illustrates how the EFT result corresponds to a UV computation.

It is possible that other UV completions could lead to cancellations that suppress FCNCs below the level expected from equation (26). However, we are not aware of models where this occurs generically, without additional fine-tuning.

C. Triangle diagram amplitudes

A new vector with purely vectorial couplings to SM fermions can only have have anomalous couplings of the form $\propto g^2 W^a W^a - g^2 B B$, since the photon and the gluon both have purely vectorial couplings as well. However, an axially-coupled vector can have anomalous couplings to $F \tilde{F}$ and $G G$. In particular, as reviewed in Appendix B, the $X_L$ coupling from fermion triangle diagrams has a fermion-mass-dependent piece proportional to the axial coupling of $X$ to the fermion. Consequently, even if the chiral fermion content of the SM + $X$ EFT is non-anomalous, the differing masses of the SM fermions...
will give enhanced $X_L$ production, unless the external momenta are small or large compared to all of the relevant fermions masses.

While the $GG$ coupling does result in enhanced production of high-$p_T$ $X$ at the LHC (resulting in monojet bounds, if $X$ registers as missing energy), the increase of the gluon parton distribution function at low energies more than compensates for the falling $(\text{energy}/m_X)^2$ enhancement. Consequently, $X$ production from gluons is actually dominated by gluons of the minimum allowed energies; if $m_X \gtrsim \Lambda_{\text{QCD}}$, there is no parametric longitudinal mode enhancement for the overall production rate.

As reviewed in section IIIE, the $FF$ coupling leads, for light enough $X$, to Primakoff process production in stars [86], giving constraints from stellar energy loss arguments.

D. Radiation from charged current decays

In models with weak-isospin breaking, the $Wq_dq_d$ or $Wl\bar{\nu}_l$ vertices may violate $U(1)_X$. This leads to charged current decays that radiate $X_L$ at an $(\text{energy}/m_X)^2$ enhanced rate. The simplest example is $W \rightarrow l\nu X$ decays [15]. While using branching ratios places weak constraints on the coupling of $X$ ($f_X \gtrsim 10\text{ GeV}$), it is possible that ‘bump-hunt’ searches in the invariant mass of the $X$ decay products could give better sensitivity.

For $Wl\bar{\nu}_l$ couplings, the $\pi^+ \rightarrow l^+\nu_lX$ decay, and the analogous charged kaon decay, are potentially sensitive channels. If $X$ decays to $e^+\nu_e^-$, then the leading competing SM decay is $\pi^+ \rightarrow e^+\nu_e(\gamma^* \rightarrow e^+e^-)$. At tree level in chiral perturbation theory, the SM decay is helicity suppressed, since the current that the pion couples to is conserved in the massless-lepton limit. In contrast, since $X$ couples differently to the electron and the neutrino, the helicity suppression is lifted in the $\pi^+ \rightarrow e^+\nu_eX$ decay, which is therefore enhanced by $m_{\pi^+}^2/(m_e^4f_X^2)$ compared to the SM decay.

Experimentally, searches for the $\pi^+ \rightarrow e^+\nu_e(h \rightarrow e^+e^-)$ decay, where $h$ is a light scalar Higgs, constrain the branching ratio of this decay to be $\lesssim 10^{-9}$ [87] for $m_h \gtrsim 10\text{ MeV}$ (for comparison, the $\pi^+ \rightarrow e^+\nu_e\gamma\gamma$ branching ratio is measured to be $3.2(\pm 0.5) \times 10^{-9}$ [87]). Taking a toy model where $X$ couples only to electrons, and not to pions or neutrinos (e.g. as occurs in the modified ‘$B - L$’ model of [8]), discussed in section IV, the $\pi^+ \rightarrow e^+\nu_eX$ branching ratio is

$$\text{Br}_{\pi^+ \rightarrow e^+\nu_eX} \approx 10^{-9} \left(\frac{200\text{ GeV}}{f_X}\right)^2.$$  

E. Non-collider constraints

At smaller masses and couplings, stellar energy loss arguments [86], and bounds on the decay of an $X$ abundance produced in the hot early universe [88], will provide additional constraints (c.f. the discussion in section II G). If longitudinal mode production is comparable to or dominates transverse production, as will be true in many cases, these constraints will be analogous to those for ALPs. For $m_X < 2m_e$, the allowed SM decays of $X$ are to neutrinos, or the (very slow) loop-induced decay to more than two photons — consequently, the bounds from late-time cosmological decays will be significantly different to the ALP case, where the two-photon decay is allowed. Additionally, if the vectorial couplings of $X$ are significantly larger than the axial ones, then the $X$ decay rate will be mainly set by the vectorial couplings, changing the relation between production and decay rates.

Stellar energy loss bounds give [89]

$$c_e^A f_X^{-1} \lesssim (10^9\text{ GeV})^{-1},$$
$$c_N^A f_X^{-1} \lesssim (10^8\text{ GeV})^{-1},$$

(31)

where $c_N^A$ is the axial coupling to nucleons — these are simply translations of the coupling bounds for an axion-like particle, applied to longitudinal $X$ emission. Bounds from horizontal branch and red giant stars, which constrain $c_A^A$ and $A_{X\gamma\gamma}$, apply at $m_X \lesssim 10\text{ MeV}$, while those from SN1987A, which constrain $c_{eX}^A$ and $A_{X\gamma\gamma}$, apply at $m_X \lesssim 100\text{ MeV}$. Supernova cooling bounds and cosmological decay constraints do not apply at large enough couplings, when the vector interacts too strongly to decay too fast (though as per section II G, there will still be cosmological constraints at small enough vector masses).

These astrophysical bounds have strong implications for searches for a fifth force mediated by an axially-coupled $X$. In particular, the spin-spin interactions induced by the exchange of $X$ [90][92] are far better constrained indirectly through (31), than they are by laboratory searches for spin-dependent forces.

In addition to constraints coming from the ALP-like behaviour of longitudinal modes, there are also constraints which are based simply on the form of the vector’s coupling to fermions. For example, if the products $c_{e\gamma\gamma}^A$ or $c_{e\nu_e}^V$ are non-zero, then $X$ exchange leads to atomic parity violation (APV), with the $c_{e\gamma\gamma}^A$ term being much more strongly constrained due to the coherent coupling to the nucleus. Measurements of the effective ‘weak charge’ of $^{133}\text{Cs}$ match the SM value to better than the percent level [93][94]. For $m_X \gtrsim 3\text{ MeV}$, the $X$ vector effectively mediates a contact-operator interaction between the nucleus and the electrons, giving a constraint of the form, [95]

$$|0.47 c_{e\gamma}^V c_{eX}^A + 0.53 c_{e\gamma}^V c_{eX}^A|^{1/2} f_X^{-1} \lesssim (10\text{ TeV})^{-1}.$$  

(32)

Also, if $X$ couples to neutrinos and electrons, then $\nu - e$ scattering measurements place strong constraints on the product of the $\nu$ and $e$ couplings. Again, for $m_X$ greater than the scattering momentum transfer, which is generally between $1-100\text{ MeV}$, these give constraints of the form $c_{e\nu} c_{eX}^{-1} \lesssim \text{TeV}^{-1}$ [96][97].
F. Comparison of constraints

To compare our new constraints to each other, and to existing bounds, we need to choose a specific model for the couplings of the new vector. A convenient choice is the $U(1)_R$ 2HDM model of [77], as discussed in section III B. This provides a UV-complete model, in which the only light new state is a vector with chiral couplings to the SM fermions. The SM + $X$ EFT within this model is anomalous, with the anomaly cancelled in the UV by heavy fermions getting their mass from a SM-singlet VEV. For simplicity, we choose the model parameters such that there is no $Z$-$X$ mass mixing — this means that the couplings to SM fermions are purely right-handed, apart from loop-induced kinetic mixing contributions.

Figure 3 illustrates a selection of the constraints we have discussed above, applied to the $U(1)_R$ model. Compared to the APV constraints, which are the most stringent in the existing literature for our parameters, the new constraints are significantly stronger across a wide mass range. In particular, the $X$ production rate at colliders and in proton beam dump experiments is significantly larger than the predictions of e.g. [77], which did not take into account $X_L$ production in meson decays. Compared to the analyses of longitudinal mode production by Fayet [14, 78], we obtain improved constraints by calculating penguin diagram contributions to flavor-changing meson decays, which are parametrically larger than other new contributions.

For $m_X > 2m_e$, the decay times for the couplings shown in the plot are $\ll 1$ sec, so the discussion of cosmological bounds in section II C applies. Supernova cooling bounds will constraints some region of couplings, though we leave computing these to future work.

The form of Figure 3 depends on the similar quark and lepton couplings in the $U(1)_R$ model. The constraints on a vector with leptophilic or leptophobic couplings will fit together differently, and some of the processes that we do not show in Figure 3, such as $\pi^\pm \to e^\pm \nu X$, can become important. Furthermore, if $X$ has dominantly invisible decays, the various missing energy searches will extend up to higher $m_X$, again resulting in strong coupling constraints, as illustrated in the right-hand panel of Figure 3.

IV. CONSTRAINTS ON MODELS

The new constraints that we have derived will generally place strong bounds on the couplings of a new light vector. In this section, we illustrate this by considering a number of models from the literature, beyond the baryon number vector and $U(1)_R$ models discussed in earlier sections.

To avoid constraints from neutrino couplings, some models introduce purely right-handed couplings to fermions. An example is the $\mu_R$ model of [8], which introduces a light vector with generation-specific couplings to right-handed muons, motivated by anomalies in low-energy muon physics. This leads to $XZ\gamma$ and $X\gamma\gamma$ anomalous couplings, and LHC $Z \to \gamma X$ searches should be able to place strong constraints on the coupling than the processes considered in [8] [48]. Furthermore, the $\mu_R - \tau_R$ version of the same model would avoid $Z \to \gamma X$ constraints, but could be probed via $e^+e^- \to \tau^+\tau^-X$ processes at $B$-factories, as discussed further in section V.

For an example featuring right-handed couplings to quarks and leptons, the model of [5] proposes a light ($10 - 100$ MeV) vector with axial couplings to first-generation fermions, which would affect the rare $\pi^0 \to e^+e^-$ decay. One point to note is that, while they only demand that the electron couplings are mostly right-handed (to suppress neutrino constraints), the down-type quark couplings should also be right-handed to avoid significant extra constraints from FCNCs, as discussed in section III B.

If we consider right-handed couplings to all fermions, then unless the couplings are chosen to cancel anomalies in the EFT, anomalous production constraints will give strong bounds. Doing this requires an electron coupling $\sim 3$ times larger than their fiducial parameters, making the model more constrained. This logical sequence illustrates some of the non-obvious requirements on light vector models.

It has recently been claimed that measurements of $^8$Be decays provide evidence for the existence of a new light vector of mass $\approx 17$ MeV, coupling to electrons and nucleons [13]. A number of papers have attempted to construct models for such a vector [21, 150, 142]. Since neither a dark photon or a $B - L$ vector can account for the data (the required couplings are excluded by other constraints, such as $\pi^0$ Dalitz-type decays and neutrino scattering), it provides an interesting case study. Here, we point out that a variety of the models in the literature are ruled out by our new constraints, unless extra fine-tuning at or above the weak scale is introduced:

- The baryon number vector model of [9] requires a coupling $g_X \gtrsim 6 \times 10^{-4}$ paired with a large kinetic mixing $\epsilon \gtrsim 10^{-3}$. These parameters result in $\text{Br}(B \to K^*X) \approx 2 \times 10^{-4}$ from the anomalous $XWW$ coupling, well above the experimental bound of $\Delta \text{Br}(B \to K^*e^+e^-) \lesssim 10^{-6}$ (see section II F). [111]

- The $B - L$ model of [9] has the SM leptons and neutrinos mix with new fields, giving the light mass eigenstates (the physical leptons and neutrinos) altered couplings to a $B - L$ boson. This allows the authors of [9] to avoid the stringent bounds coming from neutrino-electron scattering, by suppressing the coupling of the new vector to neutrinos. However, this weak-isospin breaking in the EFT means that $\pi^\pm \to e^\pm \nu X$ decays, as discussed in section III D, are energy-enhanced. In the model of [9], in which $X$ has dimension-4 couplings to electrons, but not to pions or neutrinos, the induced
FIG. 3. **Left panel:** Constraints on a vector $X$ with generation-universal couplings to right-handed quarks and leptons, assuming no additional invisible $X$ decay channels. Colored regions with solid borders indicate constraints from visible decays, dashed borders correspond to missing energy searches, the dot-dashed borders indicate displaced searches, and dotted borders denote projections based on current expected sensitivities. The gray regions indicate constraints which do not utilize the energy-enhanced production of the longitudinal mode. The new constraints come from searches for $K \rightarrow \pi X$ (green) [50, 53, 54], $B \rightarrow K X$ (blue) [15, 19], $Z \rightarrow X \gamma$ (red) [30, 51], very displaced decays at the CHARM proton beam dump experiment [58], and monojets [98]. The enhanced $K \rightarrow \pi X$ decays result in larger $X$ production than computed in naive analyses [59, 60]. The $\Upsilon$ and $J/\psi$ constraints use [80–82]. The ‘anomaly’ line shows the approximate region in which anomaly-cancelling fermions would be light enough to have been detected [27] and the $e^+e^- \rightarrow X \gamma$ constraint is from a search for dark photons [99].

**Right panel:** As above, but with the assumption that $X$ dominantly decays invisibly.

The $\pi^+ \rightarrow e^+\nu e X$ branching ratio is

$$
\text{Br}_{\pi^+ \rightarrow e^+\nu e X} \simeq 1.5 \times 10^{-9} \left( \frac{g_{eX}}{10^{-4}} \right)^2 \left( \frac{17 \text{ MeV}}{m_X} \right)^2, \tag{33}
$$

while they require $g_{eX} > 3 \times 10^{-4}$ to account for the claimed anomaly. As per section [III]D, the experimental constraint on this branching ratio, for $X \rightarrow e^+e^-$, is $\lesssim 10^{-9}$. The combination of bounds such as equation (33) and neutrino scattering constraints is particularly powerful in constraining $B - L$-based models, since they are both effects in the low-energy theory, which cannot be fine-tuned away by UV physics (in particular, neutrino-electron scattering is measured over a range of momentum transfers, so cannot be cancelled by a contact operator).

In addition, since the new fermions which mix with the SM states are heavy, integrating them out gives an anomalous EFT, with an effective $\propto \sum f_{X} (g^2 W^a \tilde{W}^a - g^2 B \tilde{B})$ ALP-like anomalous coupling. Though they envisage the new fermions having mass $\sim 100$ GeV, which is not heavy enough to be entirely unimportant in $Z$ decays or FCNC meson decays, an approximate calculation shows that their fiducial parameters should be fairly comfortably ruled out by $B \rightarrow K^* X$ decays.

- Ref. [100] proposes a light vector with a combination of (generation-universal) axial and vector couplings to quarks. At the level of nuclear amplitudes, the axial-current induced transitions are enhanced compared to vector-current induced transitions by the ratio of the proton mass to the energy of the transition. Consequently, parametrically smaller axial-vector couplings can result in the same strength nuclear transitions, and the authors of [100] find a small window of allowed couplings. The authors find that couplings of $O(10^{-5})$ can explain the excess. However, flavour-universal axial-vector couplings of this strength are in tension with our meson decay bounds, as per the results of previous sections. In particular, the coupling to right-handed down-type quarks is problematic for $\Upsilon$ decay bounds (which are not UV-dependent), while the couplings to left-handed quarks, and right-handed up-type quarks, are severely constrained by FCNC $B$ decays, in the absence of extra fine-tuning.

While constraints of these kinds will generally apply to light vector models which could explain the claimed $^8\text{Be}$ anomaly, it is possible that models with additional fine-tuning of the troublesome amplitudes (e.g., fine-tuning of FCNC amplitudes via some carefully constructed initial quark non-universality of $X$ couplings) could be found. From this perspective, it is important to test the credibility of the experimental claim [13] independently of new physics models, and, if correct, scrutinize possible nuclear physics effects that could create anomalies in the
angular distribution of the electron-positron pair \[107\].

V. FUTURE EXPERIMENTAL DIRECTIONS

As we have demonstrated, the dominant production processes for new light vectors coupled to non-conserved SM currents can be very different to those for e.g. a dark photon. These processes allowed us to place more stringent bounds on the couplings of such vectors from existing experimental data. They also suggest new classes of observations, which would allow future searches to significantly improve their discovery and exclusion potential. Here, we summarise some of the new searches and signatures which could be utilised by experimental collaborations:

- **B meson decays at B factories and LHCb**: FCNC $B \to K^{(*)}X$ meson decays are one of the most powerful probes of many light vector models. For a hadronically coupled $X$, if $m_X \gtrsim m_B$ then it can decay to $\pi^0\gamma$ or $\pi^+\pi^-\pi^0$, as well as higher-multiplicity pion or kaon final states, or nucleon-antinucleon pairs, for larger masses. These would result in peaks in the invariant mass distribution for the corresponding combination of the final state hadrons at $m_X$. Given the capability of $B$-factory experiments to resolve invariant masses with charged hadrons, $\gamma$, and $\pi^0$, such “bump hunt” analyses of $B \to K^{(*)} +$ hadronic (and electromagnetic) final states can be done with existing data from BaBar and Belle. Such searches could also be added to the Belle-II program.

$B \to K^{(*)}X$ decays in which $X$ decays leptonically can also be probed at $B$ factories. Currently, the only published analyses in the $m_{ee} \lesssim m_{\pi\pi}$ regime are for the $B \to K^+e^+e^-$ decay, since there is interest in the low-$q^2$ enhancement of the SM rate. However, for our purposes, the fact that the SM rate for $B \to Ke^+e^-$ is not enhanced at low $q^2$ is actually helpful for finding new physics contributions. In general, a fine-tuned search in the $l^+l^-$ mass spectrum in such decays would significantly tighten bounds on light, leptonically decaying $X$. In addition to $B$-factory searches, LHCb also has the ability to set constraints of this kind.

In addition to FCNC decays, leptonic decays of the form $B^\pm \to l^\pm\nu + l^\pm l^-$ occur via $B^\pm \to l^\pm\nu X$, and are enhanced if $X$ has weak-isospin violating couplings to leptons. Unlike the corresponding pion and kaon decays (section III.D), the $B^\pm \to l^\pm\nu$ decays already have a very small SM branching ratio, and it is not clear whether a new physics contribution from realistically small $X$ couplings would be visible.

- **$\tau X$ production at B and tau/charm factories**: A unique capability of $B$-factories is the possibility to study $\tau^+\tau^-$ pairs in a well-controlled environment. If $X$ has an axial coupling to the $\tau$, then the $e^+e^- \to \tau^+\tau^-X$ process gives $(m_\tau/m_X)^2$ enhanced $X_l$ emission, in analogy to the $e^+e^- \to \tau^+\tau^-S$ signatures discussed in \[105\], \[109\]. Sensitivity estimates show that BaBar and Belle searches should be able to probe $|e^\pm g_{X}/m_{X}| \lesssim (150 \text{ GeV})^{-1}$ for $200 \text{ MeV} \lesssim m_X \lesssim 4 \text{ GeV}$ \[110\], for leptonic $X$ decays.

- **Kaon decays at NA62**: The NA62 kaon experiment \[53\] at CERN will give improved bounds on charged kaon decays. These include $K^+ \to \pi^+X$ FCNC decays, which will be constrained by $K^+ \to \pi^+\ell^+\ell^-$ missing energy and $K^+ \to \pi^+l^+l^-$ searches, and $K^+ \to e^+\nu X$ decays (section III.D), constrained by $K^+ \to e^+\nu l^+l^-$ searches.

- **Rare Z and W decays at the LHC**: It is well-known that LHC searches for rare EW boson decays can provide good sensitivity to exotic states (see e.g. Refs. \[110–112\]), not least due to the extremely large number of Ws and Zs produced at the LHC. Besides the $Z \to l^+l^-l'^+l'^-$ mode, which already imposes important constraints on dark photon and $L_\mu - L_\tau$ models, both collaborations should analyse $\gamma X$ final states, with $X$ decaying to $l^+l^-$, $\pi^0\gamma$, $\pi^0\pi^+\pi^-$ etc. Such decays provide strong sensitivity to light vectors with anomalous $Z\gamma X$ couplings, such as a baryon number coupled vector, and should eventually be able to supersede LEP-derived bounds (except for the $Z \to \gamma +$ invisible channel). For $X$ with a weak-isospin-violating difference in couplings to leptons and neutrinos, $W \to l\nu X$ decays may also provide constraints. If $X$ decays leptonically, then a bump hunt in the invariant mass of the $X$ decay products could offer good sensitivity.

LHC processes other than EW boson decays may also be important in some models, and can have longitudinal emission enhanced by parametrically higher energies.

- **Displaced decays**: a number of proposed and upcoming experiments will have improved sensitivity to very displaced decays of a light vector. The SHiP proton beam dump experiment \[59\] can significantly improve sensitivity to weakly-coupled $X$ produced in $B$-meson decay processes. For lower-mass $X$s produced in kaon decays, neutrino-related experiments such as the new / planned liquid argon near detectors at Fermilab \[113\] should achieve excellent discrimination of final state products, potentially giving increased sensitivity. At higher energies, the MATHUSLA proposal to build a specialized detector capable of finding very displaced decays at the LHC \[114\] could improve bounds on light, weakly-coupled vectors, though it is not clear
whether there is viable parameter space in which longitudinal production is the dominant process.

VI. DISCUSSION

The physics point underlying this paper is that, for new vector particles with dimension-4 couplings to the SM, production of the vector’s longitudinal mode is generally enhanced by \( (E/\text{vector mass})^2 \), where \( E \) is some energy or mass scale associated with the SM process. The only way to avoid this entirely is to couple to a conserved SM current, i.e. \( B - L \) or EM. Practically, coupling to special lepton currents such as \( L_\mu - L_\tau \) also makes longitudinal production negligible, since the non-conservation of the current is controlled by the small neutrino masses. Purely right-handed couplings to first-generation quarks can also have longitudinal production suppressed by small quark masses. Energy-enhanced longitudinal mode production is a well-known effect, but as we have noted, it has been ignored in many works which consider light BSM vector particles. We have discussed a number of cases in which such energy-enhanced production leads to significantly stronger constraints than derived in existing literature, and pointed out how these can rule out a number of phenomenologically-motivated models. As well as setting constraints, our work highlights how future measurements and analyses may allow high-energy experiments to be even more sensitive probes of weakly-coupled physics.

ACKNOWLEDGMENTS

We thank Masha Baryakhtar, Emilian Dudas, Jesse Thaler, and Yue Zhao for helpful discussions. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development & Innovation. JD is supported in part by the NSF through grant PHY-1316222.

Appendix A: Constraints on new SM-chiral fermions

To cancel electroweak anomalies with new SM-chiral fermions getting their mass from the SM Higgs, we would need at least two new \( SU(2)_L \) doublets. \cite{115} considers the EWP and Higgs decay constraints on additional chiral fermions (see also \cite{116}), finding that more then two new doublets are ruled out, but that two new doublets (along with four new hypercharged singlets as their partners) are marginally allowed, for suitable choices of the hypercharge.

For these allowed hypercharge assignments, the new states are all EM charged. The only way for these charges to be integers is for the states to have charges 1 and 2. This leads to a \( \sim 10\% \) increase in the \( h \rightarrow \gamma \gamma \) width (by flipping the sign of the amplitude), and a \( \sim 130\% \) increase in the \( h \rightarrow Z \gamma \) width. With 300 fb\(^{-1}\) of LHC data, the projections of the fractional measurement accuracy for these decay channels are \( \sim 0.14 \) and \( \sim 0.45 \) respectively \cite{117}, so such deviations so be detected or ruled out (see also \cite{118} for proposed methods of improving the \( Z \gamma \) measurement). This feature is true generally — deviations will show up in either \( h \rightarrow \gamma \gamma \) or \( h \rightarrow Z \gamma \) for any choice of the hypercharge.

If we also consider direct production signals, it is likely that any such model would already have been seen at the LHC. If the new fermions have non-integer EM charges (in which case at least one has charge \( \geq 0.8 \)), then they are stable. As well as the problems arising from a possible cosmological relic population of stable non-integer-charged states, there are LHC constraints on such particles, which gives a lower mass limit of \( \sim \) 700 GeV \cite{119}.

For the case where the charges are 1 and 2, the \( Q = 2 \) fermion can decay to a \( W \) boson and the \( Q = 1 \) fermion, while the latter can mix with the SM leptons, and decay to \( Zl \) or \( hl \). \cite{120} considers the signatures of these kinds of new fermions, finding that \( \sim 100 \text{ fb}^{-1} \) of 14 TeV data should place a mass limit of \( \gtrsim 500 \text{ GeV} \). To make the new chiral fermions heavier than these lower bounds, we would need to give them very large Yukawa couplings, which introduces strong coupling problems.

In summary, the next LHC run should, if it finds no deviations from the SM, fairly robustly exclude next chiral fermions obtaining their mass from the SM Higgs, and even now, constructing models with such fermions would be a very delicate task.

Appendix B: Anomalous amplitudes

If \( \chi_i \) are massless chiral (left-handed) fermions which couple to vector particles \( V_j \), with \( \mathcal{L} \supset g_{ij} V^\mu_i \bar{\chi}_i \gamma_\mu \chi_j \), then the longitudinal amplitude for the \( V_1V_2V_3 \) triangle diagram is

\[
-(p + q)_\mu \mathcal{M}^{\mu\nu\rho} = \frac{1}{12\pi^2} \epsilon^{\rho\lambda\sigma} p_\lambda q_\sigma \sum_i g_{1i} g_{2i} g_{3i} ,
\]

\[
\mathcal{M}^{\mu\nu\rho} \equiv \sum_i V^\mu_i \mathcal{A}_i \mathcal{B}_i \mathcal{C}_i , \tag{B1}
\]

for any regularisation method that respects the symmetry between the three legs. This is termed the ‘consistent anomaly’ \cite{29}, with \( p_\mu \mathcal{M}^{\mu\nu\rho} \) and \( q_\rho \mathcal{M}^{\mu\nu\rho} \) being given by the obvious symmetrical expressions. For vectors whose couplings are non-diagonal, such as non-abelian gauge bosons, the triangle amplitude is given by the obvious modification of equation \cite{31} where there must be no overall fermion flavor change around the loop.

In many circumstances, it is helpful to use a regularisation method which does not treat the legs symme-
cally. For example, the gauge symmetries corresponding to some of the vectors may be broken, but others preserved. If we evaluate the $XBB$ triangle diagram below, using a regulator which preserves the $U(1)_B$ gauge symmetry, we obtain the ‘covariant anomaly’

$$-(p + q)_\mu M^{\mu \nu \rho} (p - q, p, q) = \frac{1}{4 \pi^2} \epsilon^{\nu \rho \lambda \sigma} p_\lambda q_\sigma \sum_i g_i X B^2_i,$$

(B2)

$$p_\nu M^{\mu \nu \rho} = 0, \quad q_\rho M^{\mu \nu \rho} = 0,$$

(B3)

$$M^{\mu \nu \rho} \equiv \sum_i X_\mu \biggrarrow p_i B_\nu \biggrarrow B_\rho .$$

In a regularisation scheme which is symmetric between $X$ and $B$, we can obtain the ‘covariant’ result by including an explicit Wess-Zumino term

$$\mathcal{L} \supset \frac{1}{6 \pi^2} \epsilon^{\mu \nu \rho \sigma} X_\mu B_\nu \partial_\rho B_\sigma \sum_i g_i X B^2_i .$$

(B4)

Within a UV completion in which both the $U(1)_X$ and $U(1)_B$ symmetries are only spontaneously broken, this WZ term will arise from integrating out heavy anomaly-canceling fermions, as illustrated in section IIC.

We can find explicit examples of these different types of regulators by considering different forms of Pauli-Villars (PV) regularisation. For the covariant anomaly, if the fermions in the loop have $B$-preserving but $X$-breaking Dirac masses, then we can regulate by introducing one heavy Dirac PV fermion for each of the Dirac fermions in the original theory. Since the couplings of the PV fermions preserve $B$, we obtain the covariant amplitude (in the limit where the original fermions have very small masses — we will return to the mass dependence of the amplitude below).

To obtain the consistent anomaly, we can introduce a PV Dirac fermion for every Weyl fermion in our original theory (rather than one per pair of original fermions), with only the left-handed parts of each PV Dirac fermion coupling to the vectors. This procedure breaks all of the gauge symmetries, unless the anomalies cancel when summing over chiral fermions. We can restore some of these symmetries, and recover the corresponding covariant anomaly, using appropriate WZ terms. This scheme has the advantage of applying unchanged in different phases of a theory, as illustrated in section IIC.

For vectors with non-diagonal couplings to the chiral fermions, as occurs with non-abelian gauge bosons, there can be four-leg (‘box’) and five-leg (‘pentagon’) anomalous amplitudes, as well as the triangle amplitudes discussed above. For the longitudinal amplitudes of an abelian vector (with diagonal couplings) against other non-abelian vectors, the pentagon diagrams are not anomalous, leaving only the triangle and box diagrams. A new vector coupling to a tree-level conserved current in the SM, as considered in section III, gives an example of the latter, with the $X - SU(2)_L - SU(2)_L$ anomaly giving $XWW$ couplings, coming from box diagrams.

1. Fermion mass dependence

Since chiral anomalies can be derived from topological considerations [29], they are independent of fermion masses. However, we are interested in longitudinal mode production, which corresponds to the divergence of the current that our vector couples to. When fermions have mass terms which break $U(1)_X$, the variation of the action under $U(1)_X$ receives contributions from these mass terms, so the divergence is the sum of these mass-dependent pieces plus the mass-independent anomaly.

For example, suppose that $X$ couples to Weyl fermions $\psi_L$ and $\psi_R$, which are the left-handed and right-handed components of a Dirac fermion with mass term $\mathcal{L} \supset -m\bar{\psi}\gamma_5\psi$ (in four-component notation). Then, we have the operator equations

$$\partial_\mu (\bar{\psi}_L \gamma^\mu \psi_L) + im\bar{\psi}_5\gamma_5\psi = \text{anomaly},$$

(B5)

$$\partial_\mu (\bar{\psi}_R \gamma^\mu \psi_R) + im\bar{\psi}_5\gamma_5\psi = \text{anomaly},$$

(B6)

where ‘anomaly’ denotes the product of field strength tensors arising from the anomaly. If

$$\mathcal{L} \supset X_\mu (g_L\bar{\psi}_L \gamma^\mu \psi_L + g_R\bar{\psi}_R \gamma^\mu \psi_R),$$

(B7)

then the matrix element between the vacuum and two gauge fields, $V_1$ and $V_2$, is

$$- (p + q)_\mu M^{\mu \rho \sigma} = \langle V_1^\sigma V_2^\rho \mid g_L \partial_\mu \bar{\psi}_L \gamma^\mu \psi_L + g_R \partial_\mu \bar{\psi}_R \gamma^\mu \psi_R \rangle .$$

(B8)

This can be rewritten using the operator equation as,

$$= (g_R - g_L)im \langle V_1^\sigma V_2^\rho \mid \bar{\psi}_5 \gamma_5 \psi \rangle + \text{anomaly} .$$

(B9)

The mass-dependent contribution is therefore set by the axial coupling of $X$ to $\psi_5$ and the vacuum-to-two-vector matrix element of the $\psi_5\gamma_5\psi$ operator. This can also be seen by direct calculation of vector triangle diagrams with massive Dirac fermions, e.g. [121]. In cases where we are considering new fermions with large, non-EW-breaking masses, such as section IIC, we are interested in e.g. $X_{BBB}$ amplitudes, where $B$ couples vectorially to the massive Dirac fermions in the loop. If $X$ couples to a given Dirac fermion $\psi$ as $\mathcal{L} \supset X_\mu \bar{\psi} (g + \gamma_5g_A)\gamma^\mu \psi$, then for a ‘consistent’ regulator we have [19, 122]

$$-(p+q)_\mu M^{\mu \nu \rho} = \frac{g_X g^2}{2 \pi^2} \epsilon^{\nu \rho \lambda \sigma} p_\lambda q_\sigma \left( \frac{1}{3} + 2m_f J_{00}(m_f, p, q) \right),$$

(B10)

where the 1/3 term is from the anomaly and the mass term gives
For large masses, the mass and the anomaly contributions combine to give the net divergence of the non-conserved current. In particular, the limit of large $m$,

$$-(p + q)_\mu M^{\mu \nu \rho} \sim -g_X g'^2 \frac{1}{3\pi^2} \epsilon^{\nu \rho \sigma \lambda} p_\lambda q_\sigma .$$  \hspace{1cm} (B12)

This is minus the contribution from the $B$-covariant WZ term. So, for a $B$-covariant regulator, the longitudinal $X$ coupling from such heavy fermions is suppressed by powers of $1/m^2$, as required by decoupling. A similar story applies for anomalous box diagrams.

As is well-known\cite{123}, the anomaly is given exactly by the 1-loop calculation (and the same applies to the matrix element of the $\bar{\psi} \gamma_5 \psi$ operator).

**Appendix C: Tree-level FCNCs**

The non-diagonal quark and neutrino Yukawa matrices in the SM mean that, even if there are no tree-level flavor-changing $X$ vertices before EWSB, they may be induced by the SM fermions getting masses. In the SM interaction basis for the quarks, we can write the Lagrangian after EWSB as

$$\mathcal{L} \supset \bar{u}_R \left( i \bar{\psi}_R \gamma^\mu p_\mu - \gamma^\mu Q_{ij} X \right) d_R \left( i \bar{\psi}_R \gamma^\mu p_\mu - \gamma^\mu Q_{ij} X \right) + \bar{d}_R \left( i \bar{\psi}_R \gamma^\mu p_\mu - \gamma^\mu Q_{ij} X \right) u_R \left( i \bar{\psi}_R \gamma^\mu p_\mu - \gamma^\mu Q_{ij} X \right) + \bar{Q}_L \left( i \bar{\psi}_R \gamma^\mu p_\mu - \gamma^\mu Q_{ij} X \right) \bar{Q}_R \left( i \bar{\psi}_R \gamma^\mu p_\mu - \gamma^\mu Q_{ij} X \right) \delta_{ij} (U_d M_d K_d^\dagger)_{ij} d_R + \bar{u}_L (U_u M_u K_u^\dagger)_{ij} u_R + h.c.) ,$$

where $M_u$ and $M_d$ are the diagonal quark mass matrices, $U_u$ and $K_u$ are unitary matrices, and $V \equiv U_d U_u$ is the CKM matrix. Without loss of generality, $Q^u$ and $Q^d$ are diagonal (after a rotation of $u_R^\dagger$ and $d_R^\dagger$). While $K_u$ and $K_d$ can be rotated away in the SM, here they have physical consequences if $Q^u$ and $Q^d$ are not the identity. However, since SM measurements are compatible with $K_u = K_d = 1$, it is possible for the $X$ couplings to right-handed fermions to be diagonal in the mass basis, without being generation-universal.

However, the CKM matrix elements can be measured in the SM, and $V \neq 1$. Thus, if we demand diagonal couplings in the mass basis, then $Q^u$ and $Q^d$ can only be proportional to each other if they are both proportional to the identity, or if one of them is zero. Hence, if there are generation-non-universal couplings to both up-type and down-type quarks, then there must be weak-isospin breaking to avoid tree-level FCNCs.

If tree-level FCNC vertices are present, the constraints on their couplings are very stringent, since amplitudes lack the loop and coupling suppressions of those discussed in section B. To give an example, if we want a vector to couple dominantly to first-generation quarks, and to have non-zero $u_L$ and $d_L$ couplings, then at least one of the $sX$ and $cX$ vertices must have coupling $\sim g_X \theta_c$, where $\theta_c \sim 0.2$ is the Cabbibo angle. The $K \to \pi X$ decay rate from the $sX$ vertex is roughly

$$\text{Br}(K \to \pi X) \sim 10^{-7} \left( \frac{g_{sX}}{10^{-10}} \right)^2 \left( \frac{100 \text{MeV}}{m_X} \right)^2$$

so experimental constraints would require that $g_X$ is extremely small.

If neutrinos are Majorana, then it is always possible that the lepton Yukawas are such that $X$ has flavor-diagonal couplings in the lepton mass basis. If they are Dirac, then a similar analysis to the above discussion for quarks applies.

\begin{thebibliography}{99}

\bibitem{1} \textit{Fundamental Physics at the Intensity Frontier} (2012) arXiv:1205.2671 [hep-ex]
\bibitem{2} R. Essig et al., in \textit{Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics}, Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013 (2013) arXiv:1311.0029 [hep-ph]
\bibitem{3} J. Alexander et al. (2016) arXiv:1608.08632 [hep-ph]
\bibitem{4} S. N. Gninenko and N. V. Krasnikov, Phys. Lett. \textbf{B513}, 119 (2001) arXiv:hep-ph/0102222 [hep-ph]
\bibitem{5} Y. Kahn, M. Schmitt, and T. M. P. Tait, Phys. Rev. \textbf{D78}, 115002 (2008) arXiv:0712.0007 [hep-ph]
\bibitem{6} M. Pospelov, Phys. Rev. \textbf{D80}, 095002 (2009)
\bibitem{7} D. Tucker-Smith and I. Yavin, Phys. Rev. \textbf{D83}, 101702 (2011) arXiv:1011.4922 [hep-ph]
\bibitem{8} B. Batell, D. McKeen, and M. Pospelov, Phys. Rev. Lett. \textbf{107}, 011803 (2011) arXiv:1103.0721 [hep-ph]
\bibitem{9} J. L. Feng, B. Fornal, I. Galon, S. Gardner, J. Smolinsky, T. M. P. Tait, and P. Tanedo, Phys. Rev. \textbf{D95}, 035017 (2017) arXiv:1608.03591 [hep-ph]
\bibitem{10} C. D. Carone and H. Murayama, Phys. Rev. Lett. \textbf{74}, 3122 (1995) arXiv:hep-ph/9411256 [hep-ph]
\bibitem{11} C. Boehm and P. Fayet, Nucl. Phys. \textbf{B683}, 219 (2004) arXiv:hep-ph/0305261 [hep-ph]
\bibitem{12} M. Pospelov, A. Ritz, and M. B. Voloshin, Phys. Lett.
