High energy spectrum of internal positrons
from radiative muon capture on nuclei

Ryan Plestid[*] and Richard J. Hill[†]

Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506, USA
and Theoretical Physics Department, Fermilab, Batavia, Illinois 60510, USA

The Mu2e and COMET collaborations will search for nucleus-catalyzed muon conversion to positrons ($\mu^- \to e^+$) as a signal of lepton number violation. A key background for this search is radiative muon capture where either (1) a real photon converts to an e^+e^- pair "externally" in surrounding material, or (2) a virtual photon mediates the production of an e^+e^- pair "internally." If the e^+ has an energy approaching the signal region then it can serve as an irreducible background. In this work we describe how the near end point internal positron spectrum can be related to the real photon spectrum from the same nucleus, which encodes all nontrivial nuclear physics.

I. INTRODUCTION

Charged lepton flavor violation (CLFV) is a smoking gun signature of physics beyond the Standard Model (SM) and is one of the most sought-after signals at the intensity frontier [1–4]. Important search channels involving the lightest two lepton generations are $\mu \to 3e$, $\mu \to e\gamma$, and nucleus-catalyzed $\mu \to e$ [1–6]. The current best limits on $\mu \to e$ come from SINDRUM-II [7,8].

The upcoming Mu2e [9] and COMET [10] experiments will either probe or discover CLFV at unprecedented levels of precision. Both experiments expect on the order of 10^{18} muon capture events, and plan to measure the ratio, $R_{e\mu}$, of $\mu \to e$ events to total muon captures, at the level of 10^{-17}.

Importantly, the Mu2e setup is "charge symmetric" such that the detection efficiencies for electrons and positrons are comparable. Thus, Mu2e will serve as a testing ground not just for the discovery of CLFV, but also for the discovery of lepton *number* violation (LNV). Specifically, on a nucleus (e.g., aluminum), the reaction

$$\mu^- + [A, Z] \to e^+ + [A, Z - 2] \qquad (1)$$

becomes a viable target for observation (see also [7]).

While neutrinoless double beta ($0\nu\beta\beta$) decay is often touted as the most promising direction for the discovery of

LNV, there do exist extensions of the SM that predict a more competitive signal in the μe sector than in the ee sector (cf. [11–13] and references therein). There also remains the looming possibility that $m_{ee} \approx 0$ (cf. Fig. 4 of [14]), rendering $0\nu\beta\beta$ insensitive to LNV if it is mediated by a light Majorana neutrino. The fact that Mu2e is charge symmetric by design yields a new handle on LNV "for free."

Unfortunately, the charge-changing nature of $\mu^- \to e^+$ can result in a substantially lower positron energy compared to the electron energy in the CLFV channel $\mu^- \to e^-$. This is driven by the mass difference between $[A, Z]$ and $[A, Z - 2]$ [15], and causes the signal region to approach, or overlap with, poorly understood SM backgrounds. The most important of these backgrounds is radiative muon capture (RMC),

$$\mu(k) + [A, Z](p) \to \nu_\mu(k') + [A, Z - 1](p') + \gamma(q), \qquad (2)$$

where, for definiteness, we assume the final state nucleus is in its ground state[1] (in the case of ^{27}Al the final state would be ^{27}Mg).

In searches for CLFV or LNV, incoming muons are the source of both the $\mu \to e$ signal and the RMC background. It is therefore critical to constrain the energy spectrum of electrons and positrons from RMC in order to discriminate signal from background.

There are two ways that RMC can contaminate the signal window in a CLFV or LNV search:

---

[*]rpl225@uky.edu
[†]richard.hill@uky.edu

[1]This assumption can be easily relaxed and a generic set of final states summed over as discussed in the Appendix D.
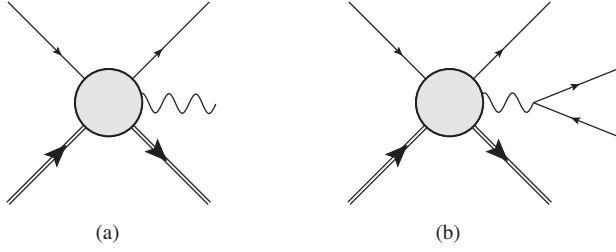
FIG. 1. Radiative muon capture on a nucleus (double lines) resulting in (a) a real photon and (b) a virtual photon that mediates the production of an electron positron pair. In this work we study how the e^+ (or e^-) spectrum in (b) is related to the photon spectrum in (a).

(1) External conversion: a (real) photon is produced and interacts with surrounding material ultimately pair producing an electron positron pair.

(2) Internal conversion: a virtual photon mediates the production of an electron positron pair.

These two possibilities are shown schematically in Fig. 1. Both cases are subject to nuclear model dependence.[2]

In the case of external conversion, nuclear model uncertainties can be circumvented at either Mu2e or COMET by directly measuring the real photon spectrum from RMC on aluminium. With this information in hand, dedicated Monte Carlo simulations (including the full detector geometry) can be used to predict the resultant electron and positron spectra. In this sense, the collaborations (Mu2e and COMET) control their own fate and can empirically constrain the external RMC backgrounds relevant for their own experiment.

Internal conversion is a more formidable challenge. Early work by Kroll and Wada [23] investigated the ratio of e^+e^- production relative to single photon production for a $1 \rightarrow 2$ process. The results of their investigation showed, unsurprisingly, that the ratio is *not* calculable without microscopic theoretical input. Rather, somewhat heuristically, they suggested that the infrared enhancement of the virtual photon favors small virtualities and argued for an approximation in terms of real photon matrix elements.

In this work we critically reexamine this problem focusing specifically on the viability of using measurements of the (real) photon spectrum to predict the internal positron (or electron) spectrum. The main conclusion of our work is that the real photon spectrum is sufficient to predict the internal spectrum of positrons *near the end point* as shown in Fig. 2. This observation implies that if the real photon spectrum is measured, then the full (internal + external) electron and positron spectra can be predicted. If the photon spectrum cannot be measured, then

---

[2]The RMC rate and spectral shape for capture on hydrogen has been studied in heavy baryon chiral perturbation theory [16,17]. Nuclear corrections substantially alter the spectral shape of RMC photons [18–21], especially near the end point [22].
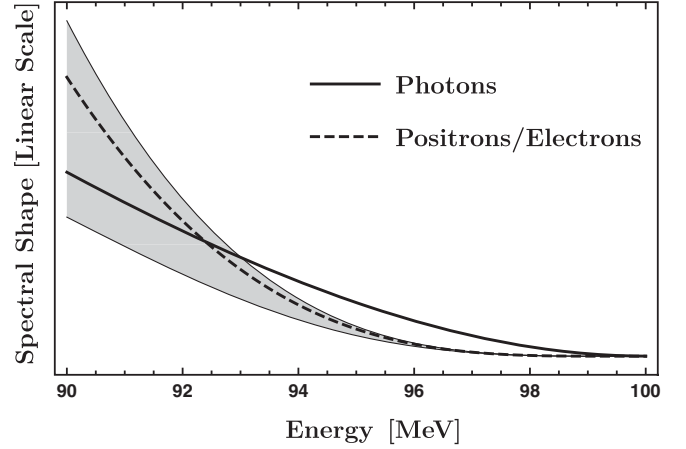

FIG. 2. Spectrum of positrons/electrons as compared to real photons near the end point assuming that the photon spectrum is given by $d\Gamma/dE_\gamma \propto (1 - 2x + 2x^2)x(1 - x)^2$ with $x = E_\gamma/q_{max}$ and $q_{max} = 100$ MeV. This corresponds to the closure approximation spectrum used in the analysis of RMC data on various nuclear isotopes in [20]. The electron/positron spectrum is computed using Eqs. (36) and (44) below. The spectrum is softened relative to photons because (1) the end point is shifted by $m_e = 0.511$ MeV, and (2) the probability of producing an e^+e^- pair rises sharply as one moves further from the end point as shown in Fig. 4. Errors due to virtual-photon nuclear matrix elements (gray band) are estimated by treating C_L and C_T (see Appendix A) as independent random variables drawn from a Gaussian distribution with unit variance.

Mu2e and COMET could use direct measurements of the total electron or positron spectrum to infer the RMC photon spectrum.

The rest of the paper is organized as follows. In Sec. II we define the nonperturbative matrix element that governs the emission of real photons and discuss how this same matrix element also governs the emission of electron positron pairs at leading order in α. Next, in Sec. III we provide an explicit formula for the real photon spectrum, which naturally leads into Sec. IV where we provide the corresponding expression for the positron spectrum. In Sec. V we discuss the near end point spectrum and demonstrate how the real photon spectrum can be used to predict internally converted positrons. Finally, in Sec. VI we summarize our conclusions, and suggest future improvements.

We also provide four Appendixes. In Appendix A we discuss how one can parametrize subleading corrections to the near end point approximation discussed above. In Appendix B we give a formal argument for some power counting details needed to justify the approximation of transversely polarized virtual photon matrix elements by their real photon counter parts. To facilitate a comparison between our work and that of Kroll and Wada we provide a short discussion of the correspondence in Appendix C. We also discuss our disagreements with their conclusions there. Finally in Appendix D we describe how to generalize the

analysis presented in the main text to the case of inclusive final states.

II. RADIATIVE MUON CAPTURE ON A NUCLEUS

The relevant S-matrix element for radiative muon capture is the overlap of an in state, $|A\mu, \text{atom}\rangle_{\text{in}}$ containing a muonic atom, with an out state $|A'\nu_\mu\gamma\rangle_{\text{out}}$ containing a recoiling nuclear system A', a muon neutrino ν_μ, and a photon γ:

$$S_\gamma = {}_{\text{out}}\langle A'\nu_\mu\gamma|A\mu, \text{atom}\rangle_{\text{in}}. \tag{3}$$

This can be expressed via the Lehmann-Symanzik-Zimmermann reduction formula as

$$S_\gamma = (2\pi)^4\delta^{(4)}(\Sigma_3 P)\epsilon^\mu {}_{\text{out}}\langle A'\nu_\mu|\hat{\mathcal{J}}_\mu|A\mu, \text{atom}\rangle_{\text{in}}, \tag{4}$$

where $\hat{\mathcal{J}}_\mu$ denotes the electromagnetic current and the notation $\Sigma_3 P$ signifies that this momentum conservation is for three on shell particles in the final state.

At leading order in α, the S matrix for internal pair production is given by (we use $\langle\hat{\mathcal{J}}_\mu\rangle \equiv {}_{\text{out}}\langle A'\nu_\mu|\hat{\mathcal{J}}_\mu|A\mu, \text{atom}\rangle_{\text{in}}$ from here on for brevity's sake)

$$S_{ee} = -(2\pi)^4\delta^{(4)}(\Sigma_4 P)\langle\hat{\mathcal{J}}_\mu\rangle_* \frac{\sqrt{4\pi\alpha}}{m_*^2}\bar{u}\gamma^\mu v. \tag{5}$$

Note that the in-out matrix element in (5) is evaluated "off shell," in that $m_*^2 \equiv (p_+ + p_-)^2 \neq 0$, where $p_\pm$ refer to the four momenta of e^+ and e^-, respectively.

Stripping off the four-momentum conserving delta function, we can therefore identify the matrix elements as

$$i\mathcal{M}_\gamma = \langle\hat{\mathcal{J}}_\mu\rangle_0\epsilon^\mu, \tag{6}$$

$$i\mathcal{M}_{ee} = -\langle\hat{\mathcal{J}}_\mu\rangle_* \frac{\sqrt{4\pi\alpha}}{m_*^2}\bar{u}\gamma^\mu v, \tag{7}$$

where in Eq. (7) we work at leading order in α. The subscript reminds us whether the matrix element has been evaluated for real, $\langle\hat{\mathcal{J}}_\mu\rangle_0$, or virtual, $\langle\hat{\mathcal{J}}_\mu\rangle_*$, photon kinematics.

To calculate the rate of decay we can make use of the standard formula[3]

$$\Gamma = \frac{1}{2M_{\text{atom}}}\int d\Phi_n\langle|\mathcal{M}|^2\rangle, \tag{8}$$

where M_{atom} is the mass of the muonic atomic (including binding energy), $\langle|\cdot|^2\rangle$ denotes an average over initial-state

---

[3]The conventional muon capture formula (see e.g., [24]) written in terms of $|\psi_{1s}(0)|^2$ can be recovered in the nonrelativistic limit by constructing the bound state $|1s, A\mu\rangle$ out of plane-wave states $|\mu(k)\rangle$ and $|A(-k)\rangle$ as described in §5 of [25].

spins and a sum over final-state spins, and $d\Phi_n$ is n-body Lorentz invariant phase space,

$$d\Phi_n = \left[\prod_{i=1}^n \frac{d^3 p_i}{2E_i(2\pi)^3}\right] \times (2\pi)^4\delta^{(4)}(\Sigma_n P). \tag{9}$$

In what follows we consider the spin-averaged matrix element which is rotationally invariant. This allows us to choose our coordinate system to lie with the $\hat{z}$ axis along the direction of the photon's momentum without loss of generality.

III. REAL PHOTON SPECTRUM

Let us introduce the tensor

$$\mathcal{J}_*^{\mu\nu} = \sum_{\text{spins}}\langle\hat{\mathcal{J}}^\mu\rangle_* \times \langle\hat{\mathcal{J}}^{\nu\dagger}\rangle_*, \tag{10}$$

where an average over initial spin states, and sum over final spin states is performed.

The real photon matrix element squared, $\langle|\mathcal{M}_\gamma|^2\rangle$, can be expressed in terms of this tensor as

$$\langle|\mathcal{M}_\gamma|^2\rangle = (-g_{\mu\nu}^\perp)\mathcal{J}_0^{\mu\nu}, \tag{11}$$

where we have used $\sum\epsilon_\mu\epsilon_\nu^* = -g_{\mu\nu}^\perp$ for the sum over physical photon polarizations. The differential rate (or spectrum) of real photons in the lab frame is then given by

$$\frac{d\Gamma_\gamma}{dE_\gamma} = \frac{1}{2M_{\text{atom}}}\int d\Phi_3[\mathcal{J}_0^{11} + \mathcal{J}_0^{22}]\delta(q_0 - E_\gamma). \tag{12}$$

IV. POSITRON SPECTRUM

The matrix element for e^+e^- creation can be expressed in terms of $\mathcal{J}_*^{\mu\nu}$ as

$$\langle|\mathcal{M}_{ee}|^2\rangle = \frac{4\pi\alpha}{m_*^4}L_{\mu\nu}\mathcal{J}_*^{\mu\nu}, \tag{13}$$

where the lepton tensor $L_{\mu\nu}$ is defined as

$$L_{\mu\nu} = \text{Tr}[(\slashed{p}_- + m_e)\gamma_\mu(\slashed{p}_+ - m_e)\gamma_\nu]. \tag{14}$$

The positron spectrum is then given by

$$\frac{d\Gamma_{ee}}{dE_+} = \frac{1}{2M_{\text{atom}}}\int d\Phi_4 \frac{4\pi\alpha}{m_*^4}L_{\mu\nu}\mathcal{J}_*^{\mu\nu}\delta(\mathfrak{E}_+ - E_+), \tag{15}$$

where $\mathfrak{E}_+$ is a function of the phase space variables that will be specified explicitly below in Eq. (26).

At this point, to make contact with the real photon spectrum, it is important to decompose phase space appropriately. This can be done via [26]

$$d\Phi_4 = d\Phi_{3*} \frac{dm_*^2}{2\pi} d\Phi_2(\gamma_* \to e^+ e^-), \qquad (16)$$

where Φ_{3*} is the three body phase space for A', ν_μ, and a particle of mass m_*. Let us compare to the massless case. We will explicitly integrate over the A' system's momentum such that only the energy conserving delta function, $\delta(\Sigma_{3*}E)$ and the neutrino-photon phase space remain

$$d\Phi_{3*} = \frac{1}{2E_{A'}} \frac{d^3 k'}{2k'(2\pi)^3} \frac{d^3 q}{2q_0(2\pi)^3} (2\pi)\delta(\Sigma_{3*}E),$$

$$= \frac{1}{2E_{A'}} \frac{d^3 k'}{2k'(2\pi)^3} \frac{d\Omega_q}{(2\pi)^3} \times \frac{\beta_* q_0}{2} dq_0 (2\pi)\delta(\Sigma_{3*}E),$$
$$(17)$$

where $\beta_* = q/q_0 = \sqrt{1 - m_*^2/q_0^2}$, and we have made use of $dq/dq_0 = 1/\beta_*$. This is to be compared with the case of a massless photon which does not have the factor of β_*

$$d\Phi_3 = \frac{1}{2E_{A'}} \frac{d^3 k'}{2k'(2\pi)^3} \frac{d\Omega_q}{(2\pi)^3} \times \frac{q_0}{2} dq_0 (2\pi)\delta(\Sigma_3 E). \quad (18)$$

The energy conservation condition is (for a recoiling A' of mass M)

$$\sum_{3*} E = k + q_0 + \underbrace{\sqrt{k^2 + |\vec{q}|^2 + 2|\vec{q}|k\cos\theta_{\nu\gamma} + M^2}}_{E_{A'}} - M_{\text{atom}},$$
$$(19)$$

where $|\vec{q}|^2 = q_0^2 - m_*^2$ and $\theta_{\nu\gamma}$ is the opening angle between the photon and neutrino in the lab frame. Integrating over the neutrino energy introduces a factor of $[\frac{\delta\Sigma_{3*}E}{\delta k}]^{-1}$:

$$\frac{1}{1 + \frac{k + |\vec{q}|\cos\theta}{E_{A'}}} = \frac{E_A'}{M_{\text{atom}} - (1 - \beta_* \cos\theta_{\nu\gamma})q_0}. \quad (20)$$

The factor of $E_{A'}$ cancels against the factor of $E_{A'}$ in the denominator of Eqs. (17) and (18) such that the phase space for a massive photon can be related to the phase space for a massless photon via

$$d\Phi_{3*} = d\Phi_3 \times \beta_* F_*, \qquad (21)$$

with Φ_3 being independent of m_* and $d\Phi_2$, and the variable F_* being given by [cf. Eq. (2) of [24]]

$$F_* = \left[\left(\frac{\delta\Sigma_3 E}{\delta k}\right) / \left(\frac{\delta\Sigma_{3*}E}{\delta k}\right) \right] \times \frac{k'_*}{k'},$$

$$= \frac{M_{\text{atom}} - (1 - \cos\theta_{\nu\gamma})q_0}{M_{\text{atom}} - (1 - \beta_* \cos\theta_{\nu\gamma})q_0} \times \frac{k'_*}{k'}, \qquad (22)$$

where the various energetic factors in the phase space measure are modified so that they satisfy Eq. (19).

In terms of these variables we can then write

$$\int d\Phi_4 = \int d\Phi_3 \int_{4m_e^2}^{q_0^2} \frac{dm_*^2}{2\pi} \beta_* F_* \int d\Phi_2, \qquad (23)$$

with Φ_2 the two-body phase space for a virtual photon of mass m_* decaying into an $e^+ e^-$ pair. We have inserted the integral symbols explicitly to emphasize the order in which the they must be performed.

Let us now study Eq. (15), performing the integrations sequentially from right to left as suggested by Eq. (23). Since the decomposed phase space is factorized into independently Lorentz invariant pieces we can carry out the integration in the frame of our choice. The quantity $\mathcal{J}_{\mu\nu}$ is independent of the two-body phase space Φ_2 and so

$$\int d\Phi_2 L_{\mu\nu} \delta(\mathfrak{E}_+ - E_+), \qquad (24)$$

in the rest frame of the photon. Two body phase space in this frame is given by

$$d\Phi_2 = \frac{d\cos\vartheta d\varphi}{32\pi^2} \beta_e, \qquad (25)$$

where $\beta_e = \sqrt{1 - 4m_e^2/m_*^2}$ and coordinates are defined such that $\hat{z}$ is parallel to the boost direction that connects the photon's rest frame to the lab frame. As already discussed above, this choice can be made without loss of generality because of the spherical symmetry of the spin-averaged matrix element $\langle |\mathcal{M}|^2 \rangle$.

In terms of these variables the function $\mathfrak{E}_+$ that determines the positron's energy is given by

$$\mathfrak{E}_+(q_0, \cos\vartheta) = \frac{q_0}{2}(1 + \beta_e \beta_* \cos\vartheta), \qquad (26)$$

where $\beta_* = \sqrt{1 - m_*^2/q_0^2}$ is (as above) the velocity of the virtual photon.

The delta function is independent of φ and so it is convenient to define $\langle L_{\mu\nu} \rangle_\varphi = \int \frac{d\varphi}{2\pi} L_{\mu\nu}$ given explicitly by

$$\langle L_{\mu\nu} \rangle_\varphi = m_*^2 \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 - (1 - \cos^2\vartheta)\beta_e^2 & 0 & 0 \\ 0 & 0 & 2 - (1 - \cos^2\vartheta)\beta_e^2 & 0 \\ 0 & 0 & 0 & 2(1 - \cos^2\vartheta\beta_e^2) \end{pmatrix} \quad \text{(in photon rest frame).} \qquad (27)$$

Then we find

$$\int d\Phi_2 L_{\mu\nu}\delta(\mathfrak{E}_+ - E_+) = \frac{\beta_e}{8\pi}\frac{1}{q_0\beta_e\beta_*}\langle L_{\mu\nu}\rangle_\varphi|_{\cos\theta}, \quad (28)$$

where we have used (note the different fonts of ϑ vs θ)

$$\delta(\mathfrak{E}_+ - E_+) = \frac{2}{q_0\beta_*\beta_e}\delta(\cos\vartheta - \cos\theta). \quad (29)$$

The emission angle of the positron, $\cos\theta$, is found by solving $E_+ = \mathfrak{E}_+(q_0, \cos\theta)$, being given explicitly by

$$\cos\theta(E_+, q_0, m_*) = \frac{2E_+/q_0 - 1}{\beta_*\beta_e}. \quad (30)$$

Obviously, $\cos\theta$ must lie in the interval $[-1, 1]$ and so Heaviside functions enforcing this condition should be included:

$$
\begin{aligned}
\frac{d\Gamma_{ee}}{dE_+} &= \frac{1}{2M_{\text{atom}}}\int d\Phi_3\frac{dm_*^2}{2\pi}\times\beta_*F_*\times\frac{4\pi\alpha}{m_*^4} \\
&\quad \times\frac{1}{8\pi}\frac{1}{q_0\beta_*}\langle L_{\mu\nu}\rangle_\varphi\mathcal{J}_*^{\mu\nu}\Theta(1 - |\cos\theta|), \\
&= \frac{1}{2M_{\text{atom}}}\int d\Phi_3\frac{\alpha}{4\pi q_0}\int dm_*^2 F_*\left[\frac{1}{m_*^4}\langle L_{\mu\nu}\rangle_\varphi\mathcal{J}_*^{\mu\nu}\right] \\
&\quad \Theta(1 - |\cos\theta|), \\
&= \frac{1}{2M_{\text{atom}}}\int d\Phi_3\frac{\alpha}{4\pi q_0}\int dm_*^2\frac{F_*}{m_*^2} \\
&\quad \times[(2 - [1 - \cos^2\theta]\beta_e^2)(\mathcal{J}_*^{11} + \mathcal{J}_*^{22}) \\
&\quad + 2(1 - \cos^2\theta\beta_e^2)\mathcal{J}_*^{33}]\Theta(1 - |\cos\theta|). \quad (31)
\end{aligned}
$$

In Eq. (31) we have contracted $\mathcal{J}_*^{\mu\nu}$ against the lepton tensor evaluated in the rest frame of the photon.[4]

Now let us define auxiliary variables to condense the notation:

$$T_*^2 = \frac{[\mathcal{J}_*^{11} + \mathcal{J}_*^{22}] - [\mathcal{J}_0^{11} + \mathcal{J}_0^{22}]}{\mathcal{J}_0^{11} + \mathcal{J}_0^{22}}, \quad (32)$$

$$L_*^2 = \frac{\mathcal{J}_*^{33}}{\mathcal{J}_0^{11} + \mathcal{J}_0^{22}}. \quad (33)$$

The quantities T_*^2, and L_*^2 are related to the transverse and longitudinal matrix elements, respectively. Notice that $T_*^2(m_* = 0) = 0$ by construction such that $T_*^2 \sim O(m_*^2)$ in the small m_* limit.

Having introduced all of the necessary ingredients, we can now write the positron spectrum as if it were produced by real photons converting "internally" via a probabilistic process described by a function P_{int}

$$\frac{d\Gamma_{ee}}{dE_+} = \frac{1}{2M_{\text{atom}}}\int d\Phi_3[\mathcal{J}_0^{11} + \mathcal{J}_0^{22}]P_{\text{int}}(E_+|q_0, \Pi). \quad (34)$$

The symbol Π represents a set of 3-body phase space variables (e.g., the photon-neutrino opening angle $\cos\theta_{\nu\gamma}$), and we have introduced the internal conversion probability

$$
\begin{aligned}
P_{\text{int}}(E_+|q_0, \Pi) &= \frac{1}{q_0}\frac{\alpha}{4\pi}\int\frac{dm_*^2}{m_*^2}F_* \\
&\quad \times [(2 - [1 - \cos^2\theta]\beta_e^2)(1 + T_*^2) \\
&\quad + 2(1 - \cos^2\theta\beta_e^2)L_*^2]\Theta(1 - |\cos\theta|), \quad (35)
\end{aligned}
$$

where $dq_0 \times \int P_{\text{int}}(E_+|q_0, \Pi)d\Pi$ can be interpreted as the probability for photons with energy between q_0 and $q_0 + dq_0$ to produce an electron positron pair.

The functions T_*^2 and L_*^2 are unconstrained by measurements of the on shell photon spectrum. As discussed above, $T_*^2 \sim O(m_*^2)$ for small virtualities, and so if $m_* \ll q_0$ then we can expect it to be small. In contrast, L_* is expected to be $O(1)$ in the $m_* \to 0$ limit.[5] Notice, however, that as $\cos\theta \to 1$ (or equivalently as $E_+ \to q_0 - m_e$) that the longitudinal matrix elements are suppressed. We will discuss this in more detail in Sec. V.

Our definition of P_{int} still depends on the unconstrained matrix elements L_*^2 and T_*^2. As we will show in the next section there exists a limit where L_*^2 and T_*^2 can be neglected. In this case Eq. (34) can be rewritten by carrying out the integration over all of the 3-body phase space except for dq_0 in which case we find

$$\frac{d\Gamma_{ee}}{dE_+} = \int dE_\gamma\frac{d\Gamma_\gamma}{dE_\gamma}P_{\text{int}}(E_+|E_\gamma), \quad (36)$$

which allows us to construct the positron spectrum using the measured photon spectrum from the same nucleus, and the calculable function $P_{\text{int}}(E_+|E_\gamma)$. We now turn our attention to the aforementioned limit in which P_{int} is independent of T_*^2 and L_*^2.

V. NEAR END POINT SPECTRUM

As discussed in the introduction, for $\mu \to e$ searches it is the high-energy tail of the RMC spectrum that is most important. This motivates studying the internal conversion

---

[4]If the longitudinal matrix element is evaluated in the lab frame then one finds $(m_*/q_0) \times [L_*^2]_{\text{lab}}$ [23]. We can relate $[L_*^2]_{\text{lab}}$ to $[L_*^2]_{\text{rest}}$ by boosting between frames, and one finds that $[L_*^2]_{\text{lab}} = (q_0/m_*) \times [L_*^2]_{\text{rest}}$ such that the factor of m_*/q_0 is precisely canceled.

[5]This is qualitatively different than the conclusion of [23]. We discuss the origin of these differences in Appendix C.

probability in the limit where $E_+ \to E_+^{(\mathrm{max})}$. This is conveniently parametrized by the dimensionless ratio formed from the electron's kinetic energy to the virtual photon energy,

$$\delta = \frac{q_0 - (E_+ + m_e)}{q_0} = \frac{E_- - m_e}{q_0}, \tag{37}$$

which tends to zero as the positron approaches its end point.

Let us now study the integral over m_* in Eq. (35) more carefully. Having restricted ourselves to a specific value of E_+, the limits of integration on m_* are supplied by the Heaviside function. This can be understood by plotting $\mathfrak{E}_+$ for $\cos\theta = \pm 1$, as a function of m_* at different fixed values of q_0 as depicted in Fig. 3. At a fixed positron energy E_+, the limits of integration over m_* can be found by solving

$$\mathfrak{E}_+(m_*, q_0, \cos\vartheta = 1) = E_+. \tag{38}$$

These are given explicitly by

$$m_*^{(\pm)} = \sqrt{2\left[m_e^2 + E_+(q_0 - E_+) \pm \sqrt{(E_+^2 - m_e^2)([q_0 - E_+]^2 - m_e^2)}\right]}. \tag{39}$$

Notice that $m_*^{(-)} \geq 2m_e$ and consequently $m_e/m_* \leq 1$ over the full range of integration. If we introduce the small parameter

$$\epsilon = \frac{m_e}{q_0}, \tag{40}$$

then the maximal value of m_* is set by

$$m_*^{(+)} = q_0[2\sqrt{(1-\delta)\delta} + O(\epsilon)]. \tag{41}$$

We therefore find that $m_*/q_0 \leq O(\sqrt{\delta})$ in the small δ limit (provided $\epsilon \ll \delta$). Similarly, for a high-energy positron to emerge from a virtual photon, it must be emitted collinear to the photon's momentum, $\cos\theta \approx 1$. Explicitly,

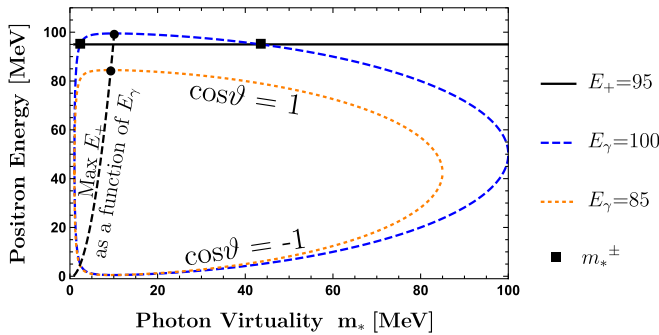$$1 - \cos\theta = 1 - \frac{2E_+/q_0 - 1}{\beta_*\beta_e} = O(\delta), \tag{42}$$

FIG. 3. Phase space for radiative muon capture. Varying $\cos\vartheta$ from -1 to 1 at fixed m_* moves from the bottom of the hull to the top of the hull. Changing $E_\gamma = q_0$ changes the size of the hull to be integrated over (i.e., orange dotted curve [$E_\gamma = 85$ MeV] $\to$ blue dashed curve [$E_\gamma = 100$ MeV]). If $E_+ \to q_0$ then only the top tip of the hull contributes to the phase space integration. Units of E_+ and m_* are in MeV.

where we treat $\beta_e \approx \beta_* \approx 1$ (see Appendix A for details). Thus, we can reexpress the internal conversion probability for $E_+ \to q_0$ as

$$P_{\mathrm{int}} = \frac{1}{q_0}\frac{\alpha}{\pi}\int_{m_*^{(-)}}^{m_*^{(+)}} \frac{\mathrm{d}m_*^2}{m_*^2} F_*$$
$$\times [2 + 2T_*^2 - O(\delta) + O(\delta) \times L_*^2],$$
$$= P_{\mathrm{int}}^{(0)}(E_+|q_0) + O(\delta). \tag{43}$$

where we have counted $T_*^2 \sim O(\delta)$ (see Appendix B for more details). The longitudinal matrix elements, by contrast, are taken to be $O(1)$ in the $m_* \to 0$ limit, but are suppressed by $(\cos^2\theta\beta_e^2 - 1) = O(\delta)$. Therefore at leading order in δ we find, setting $F_* \approx 1$ as discussed in Appendix A, that

$$P_{\mathrm{int}}^{(0)}(E_+|q_0) = \frac{1}{q_0}\frac{\alpha}{4\pi}\int_{m_*^{(-)}}^{m_*^{(+)}} \frac{\mathrm{d}m_*^2}{m_*^2} \times 2,$$
$$= \frac{1}{q_0}\frac{\alpha}{\pi}\log(m_*^{(+)}/m_*^{(-)}). \tag{44}$$

This function is plotted, along with error estimates discussed in Appendix A, in Fig. 4. The small-δ form of $m_*^{(+)}/m_*^{(-)}$ depends on the relative size of ϵ and δ. Explicitly the various scaling limits are given by (always taking $\delta, \epsilon \ll 1$)

$$\frac{m_*^{(+)}}{m_*^{(-)}} \sim \begin{cases} 2\delta/\epsilon & \delta \gg \epsilon \\ \left(\dfrac{\delta + \epsilon + \delta\sqrt{1 + 2\epsilon/\delta}}{\delta + \epsilon - \delta\sqrt{1 + 2\epsilon/\delta}}\right)^{1/2} & \delta \sim \epsilon \\ 1 + \sqrt{2\delta/\epsilon} & \delta \ll \epsilon \end{cases} \tag{45}$$

If one wishes to straddle these various regimes then the exact expressions for $m_*^{(\pm)}$, as presented in Eq. (39), should be used. This is important for applications at Mu2e and
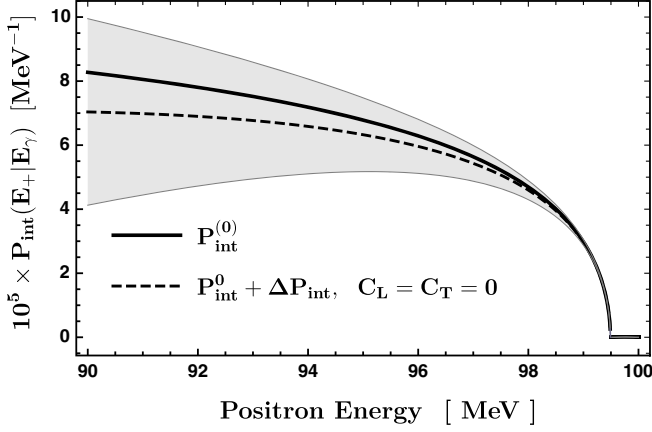
FIG. 4. Probability per differential bin (i.e., of width dE_γ) of producing a positron with energy E_+ given a photon energy of $E_\gamma = 100$ MeV, Eqs. (39) and (44). We plot $P_{\text{int}}^{(0)}$ of Eq. (43) as well as error estimates (gray band) obtained by considering an ensemble of the two parameters C_L, and C_T, that capture nuclear structure dependence (see discussion in Appendix A). We treat C_L and C_T as random variables drawn from independent Gaussian distributions such that $\langle C_L^2 \rangle = \langle C_T^2 \rangle = 1$.

COMET where energy resolution is of $O(100 \text{ keV})$ [9] such that electron mass effects can be resolved.

VI. CONCLUSIONS AND OUTLOOK

We have shown that given the photon spectrum from RMC on a nucleus, the spectrum of high energy positrons or electrons can be computed accurately with errors controlled by $\delta = (E_- - m_e)/q_0$. This allows for a robust characterization of the internal RMC positron and electron spectra near the end point. External positrons, stemming from real photons pair producing in surrounding detector material, can be calculated from the real-photon spectrum by dedicated Monte Carlo simulations that include the full detector geometry. Thus, with this present work, a measurement of the real-photon spectrum becomes entirely predictive for the purposes of determining the high-energy positron spectrum. Most importantly, this allows for nuclear physics effects, which have large theoretical uncertainties, to be included empirically with measured data.

While our study has emphasized the near end point positron spectrum, it applies equally well to the near end point electron spectrum. The reason that electrons and positrons can be treated on an equal footing is that we have *not* included the influence of the nucleus' strong Coulomb field, which, in reality, will influence the outgoing electron and positron. Near the end point, for internal conversion, this approximate neglect of Coulomb corrections may be insufficient. A high-energy positron necessarily implies a low-energy electron whose outgoing wave function is then substantially distorted by the Coulomb field of the nucleus (as is well known in the theory of beta decay [27]). This has

been discussed in a more general setting in [28], and we will study this issue in the context of the RMC end point in future work [29]. We anticipate that our formalism of an internal conversion probability can be readily adapted to account for Coulomb distortion effects.

APPENDIX A: ERROR ESTIMATES

Here we consider corrections to the universal conversion probability P_{int}. As emphasized in Eq. (45), these corrections depend on the relative size of ϵ and δ. In what follows we will consider the cases of $\epsilon \ll \delta$ [i.e., a relativistic electron satisfying $2m_e \ll (E_\gamma - E_+)$], $\epsilon \sim \delta$ [i.e., a quasirelativistic electron satisfying $2m_e \sim (E_\gamma - E_+)$], and $\epsilon \gg \delta$ [i.e., a nonrelativstic electron $2m_e \gg (E_\gamma - E_+)$] separately.

For this section it is convenient to reexpress Eq. (30) as

$$\cos\theta = \frac{2(1 - \delta - \epsilon) - 1}{\beta_* \beta_e} = \frac{1 - 2\delta - 2\epsilon}{\beta_* \beta_e}, \quad (A1)$$

where we remind the reader that $\beta_* = \sqrt{1 - m_*^2/q_0^2}$ and $\beta_e = \sqrt{1 - 4m_e^2/m_*^2}$. The ratio of $m_*^{(\pm)}$ to the photon energy q_0 will also be useful, and this can written in terms of β and ϵ as

$$\frac{m_*^{(\pm)}}{q_0}$$
$$= \sqrt{2[\epsilon + \delta(1 - \delta - 2\epsilon) \pm \sqrt{\delta(1 - \delta)(1 - \delta - 2\epsilon)(\delta + 2\epsilon)}]}.$$
$$(A2)$$

Subleading corrections to P_{int} depend on nuclear matrix elements and are therefore model dependent. We can,

however, parametrize this model dependence at next to leading order by defining

$$C_T(\Pi) = q_0^2 \left[\frac{1}{2} \frac{\delta^2}{\delta m_*^2} T_*^2 \right]_{m_*=0}, \qquad (A3)$$

$$C_L(\Pi) = [L_*^2]_{m_*=0}, \qquad (A4)$$

where C_T and C_L depend on the other phase-space variables Π ($\Pi = \cos\theta_{\nu\gamma}$, the opening angle between the photon and neutrino, for RMC). In estimating errors in the main text we treat C_L and C_T as constants independent of $\cos\theta_{\nu\gamma}$. The small-m_* behavior of the functions T_*^2 and L_*^2 can be expressed in terms of C_T and C_L via

$$T_*^2 = C_T(\Pi) \times \left(\frac{m_*}{q_0} \right)^2 + O(m_*^4), \qquad (A5)$$

$$L_*^2 = C_L(\Pi) + O(m_*^2). \qquad (A6)$$

Depending on the relative sizes of $m_*^{(\pm)}$, either of β_e or β_* can deviate from unity at next to leading order. We therefore expand both of them in a Taylor series:

$$\beta_e = \sqrt{1 - \frac{4m_e^2}{m_*^2}} \sim 1 - \frac{2m_e^2}{m_*^2}, \qquad (A7)$$

$$\beta_* = \sqrt{1 - \frac{m_*^2}{q_0^2}} \sim 1 - \frac{m_*^2}{2q_0^2}. \qquad (A8)$$

While the expansion in m_e^2/m_*^2 may look ill-behaved for small m_*, recall that $m_* \geq 2m_e$, and for $\delta \ll 1$ we always have that $m_e/m_* \lesssim O(\sqrt{\delta})$. Throughout the full region of phase space we find that $F_* \sim 1 + O(\frac{m_*^2}{q_0 M_{\mathrm{atom}}})$ and for medium-heavy nuclei these deviations from unity are at the permille level since $m_* \sim 10$ MeV, $q_0 \sim 100$ MeV, and $M_{\mathrm{atom}} \sim 30$ GeV; we therefore set $F_* = 1$ hereafter.

We may then expand Eq. (35) under the integral sign subtracting off the leading-order expression. We therefore define

$$\Delta P_{\mathrm{int}} = P_{\mathrm{int}} - P_{\mathrm{int}}^{(0)}, \qquad (A9)$$

given explicitly by

$$\Delta P_{\mathrm{int}} = \frac{\alpha}{4\pi} \int \frac{dm_*^2}{m_*^2} \left[2C_T(\Pi) \frac{m_*^2}{q_0^2} - \left(4(\delta+\epsilon) - \frac{m_*^2}{q_0^2} - \frac{4m_e^2}{m_*^2} \right) + 2 \left(4(\delta+\epsilon) - \frac{m_*^2}{q_0^2} \right) C_L(\Pi) \right],$$

$$= \frac{1}{q_0} \frac{\alpha}{\pi} \int \frac{dm_*}{m_*} \left\{ 2[\delta+\epsilon][2C_L(\Pi) - 1] + \left[C_T(\Pi) - C_L(\Pi) + \frac{1}{2} \right] \frac{m_*^2}{q_0^2} + \frac{2m_e^2}{m_*^2} \right\}. \qquad (A10)$$

In deriving the above equation we have made use of

$$(1 - \cos^2\theta)\beta_e^2 = \frac{\beta_e^2 \beta_*^2 - (1 - 2\delta - 2\epsilon)^2}{\beta_*^2} \sim 4(\delta+\epsilon) - \frac{m_*^2}{q_0^2} - \frac{4m_e^2}{m_*^2} + O(\delta^2), \qquad (A11)$$

$$(1 - \cos^2\theta\beta_e^2) = \frac{\beta_*^2 - (1 - 2\delta - 2\epsilon)^2}{\beta_*^2} \sim 4(\delta+\epsilon) - \frac{m_*^2}{q_0^2} + O(\delta^2). \qquad (A12)$$

If we integrate over m_* we then find

$$\Delta P_{\mathrm{int}} = \frac{1}{q_0} \frac{\alpha}{\pi} \left\{ 2[\delta+\epsilon][2C_L(\Pi) - 1] \log(m_*^{(+)}/m_*^{(-)}) + \left[\frac{1}{2} C_T(\Pi) - \frac{1}{2} C_L(\Pi) + \frac{1}{4} \right] \frac{[m_*^{(+)}]^2 - [m_*^{(-)}]^2}{q_0^2} \right.$$

$$\left. + \left[\left(\frac{m_e}{m_*^{(-)}} \right)^2 - \left(\frac{m_e}{m_*^{(+)}} \right)^2 \right] \right\}. \qquad (A13)$$

Different terms in Eq. (A13) will be relevant or negligible depending on whether $\epsilon \ll \delta$, $\epsilon \sim \delta$, or $\epsilon \gg \delta$. It is therefore useful to consider these limits separately.

1. Relativistic electron: $\epsilon \ll \delta$

In this section we will power count with $\epsilon \sim O(\delta^2)$ such that

$$\frac{m_*^{(+)}}{q_0} \sim 2\sqrt{\delta}, \qquad (A14)$$

$$\frac{m_*^{(-)}}{q_0} \sim \frac{\epsilon}{\sqrt{\delta}} \sim O(\delta^{3/2}). \qquad (A15)$$

We then find

$$\Delta P_{\text{int}} = \frac{2\delta}{q_0}\frac{\alpha}{\pi}\{[2C_L(\Pi) - 1]\log(2\delta/\epsilon)$$
$$+ [C_T(\Pi) - C_L(\Pi) + 1]\}, \tag{A16}$$

where all contributions proportional to ϵ have been dropped because of the $\epsilon \sim \delta^2$ power counting. The final term in Eq. (A16) receives a contribution from $[m_e/m_*^{(-)}]^2 \sim O(\delta)$ and $[m_*^{(+)}/q_0]^2 \sim O(\delta)$.

2. Quasi-relativistic electron: $\epsilon \sim \delta$

In this section we treat $\epsilon \sim O(\delta)$ in our power counting, and find

$$\frac{m_*^{(\pm)}}{q_0} \sim \sqrt{2\delta} \times \left[1 + \frac{\epsilon}{\delta} \pm \sqrt{1 + \frac{2\epsilon}{\delta}}\right]^{1/2}. \tag{A17}$$

Notice that both $m_*^{(+)}$ and $m_*^{(-)}$ are parametrically of the same size, in contrast to the case where $\epsilon \ll \delta$ where we instead found that $m_*^{(+)} \gg m_*^{(-)}$. We then find

$$\Delta P_{\text{int}} = \frac{2\delta}{q_0}\frac{\alpha}{\pi}\left\{[2C_L(\Pi) - 1]\left(1 + \frac{\epsilon}{\delta}\right)\log(m_*^{(+)}/m_*^{(-)})\right.$$
$$\left. + \sqrt{1 + \frac{2\epsilon}{\delta}}[C_T(\Pi) - C_L(\Pi) + 1]\right\}, \tag{A18}$$

where, for compactness, we have left the dependence of $m_*^{(\pm)}$ on ϵ and δ implicit inside the logarithm. The final term in Eq. (A20) receives contributions from both roots of $m_*^{(\pm)}$.

3. Nonrelativistic electron: $\epsilon \gg \delta$

In this section $\epsilon \sim O(\delta^{1/2})$ in our power counting, and

$$\frac{m_*^{(\pm)}}{q_0} \sim \sqrt{2\epsilon} \pm \sqrt{\delta}. \tag{A19}$$

Much like in the previous section both roots are of roughly the same size. We then find that

$$\Delta P_{\text{int}} = \frac{\sqrt{8\delta\epsilon}}{q_0}\frac{\alpha}{\pi}[C_T(\Pi) + C_L(\Pi)], \tag{A20}$$

where we have Taylor expanded $\log(m_*^{(+)}/m_*^{(-)})$ taking advantage of the fact that $\delta \ll \epsilon$. In this regime the error scales parametrically like $O(\sqrt{\epsilon\delta})$. This should be compared to $P_{\text{int}}^{(0)}$ in the same regime which is $O(\sqrt{\delta/\epsilon})$. The relative error is therefore $O(\epsilon)$. This stems from the photon virtuality m_* tending towards $m_*^2 \to 2m_e q_0$ in the limit that $\delta \to 0$ such that $m_*^2/q_0^2 \sim O(m_e/q_0) \sim O(\epsilon)$.

APPENDIX B: VIRTUALITY DEPENDENCE OF THE TRANSVERSE ELECTROMAGNETIC CURRENT

In Eq. (43) we have assumed that corrections stemming from L_*^2 and T_*^2 are $O(\delta)$ or smaller. The longitudinal matrix elements, L_*^2, are suppressed explicitly by a factor of $(1 - \beta_e^2 \cos^2 \theta) \sim O(\delta)$. In contrast, the transverse matrix elements, T_*^2, are not suppressed by explicitly $O(\delta)$ prefactors. For corrections from T_*^2 to be small, we therefore require that terms proportional to the virtual photon mass m_*^2 are suppressed by energetic scales of $O(q_0^2) \sim O(m_\mu^2)$ such that $T_*^2 \sim m_*^2/q_0^2 \sim O(\delta)$. One may be concerned that energy-level splittings (on the order of $\Delta E \sim$ few MeV) from low-lying nuclear excited states could enhance m_* dependent contributions leading to $O(m_*^2/\Delta E^2)$ corrections. In this section we study the structure of $\mathcal{J}_*^{\mu\nu}$ by inserting a complete set of nuclear states, and deriving a time-ordered expression where energy-splitting denominators appear explicitly.

We will study the matrix element

$$_{\text{out}}\langle \nu, A' | \hat{\mathcal{J}}_\mu | A\mu, \text{atom}\rangle_{\text{in}}$$
$$\approx \int \frac{d^3k}{(2\pi)^3} \tilde{\psi}_\mu(k)_{\text{out}}\langle \nu, A' | \hat{\mathcal{J}}_\mu | A(-k), \mu(k)\rangle_{\text{in}}$$
$$\approx \psi_\mu(0)_{\text{out}}\langle \nu, A' | \hat{\mathcal{J}}_\mu | A, \mu\rangle_{\text{in}}, \tag{B1}$$

where we have assumed a nonrelativistic treatment of the muon-nucleus bound state in the first approximation, and approximated the plane-wave matrix element by its value at $k = 0$ in the second approximation; we have also used $\psi_\mu(0) = \int \tilde{\psi}_\mu(k)d^3k/(2\pi)^3$.

We are interested in this matrix element for real- and virtual-photon kinematics at fixed photon energy q_0. We take the initial state as having zero three-momentum and an energy (i.e., rest mass) of M_{atom}. The outgoing momentum of A' is fixed by momentum conservation, and the energy of the neutrino by energy conservation. In the limit where $M_{A'} \sim M_{\text{atom}} \to \infty$ (i.e., neglecting the recoil energy of A') we find

$$E_\nu = (M_{\text{atom}} - M_{A'}) - q_0 + O(q_0^2/M_{A'}),$$
$$\approx m_\mu - q_0, \tag{B2}$$

where in the approximation we have neglected the binding energy of the muon and the mass difference $M_{A'} - M_{\text{atom}}$.

Let us evaluate $_{\text{out}}\langle \hat{\mathcal{J}}_\mu\rangle_{\text{in}}$ at leading order in G_F using Dyson's formula, such that the lepton and nuclear matrix elements can be fully separated

$$_{\text{out}}\langle\nu, A'|\hat{\mathcal{J}}_\mu|A, \mu\rangle_{\text{in}} = \langle\nu, A'|T\{e^{-i\int d^4 y \mathcal{H}_{\text{int}}(y)}\hat{\mathcal{J}}_\mu(0)\}|\mu, A\rangle,$$

$$= \frac{G_F}{\sqrt{2}}i\int d^4 y\langle\nu, A'|T\{\bar{\nu}(y)\gamma^\nu(1-\gamma_5)\ell(y)\hat{J}_\nu(y)\hat{\mathcal{J}}_\mu(0)\}|\mu, A\rangle + O(G_F^2), \quad (B3)$$

where $\hat{J}_\nu$ is the weak-hadronic current. Furthermore, the electromagnetic current is a linear combination of a leptonic and hadronic piece

$$\hat{\mathcal{J}}_\mu = \hat{\mathcal{H}}_\mu + \sum_\ell - e\bar{\ell}\gamma_\mu\ell. \quad (B4)$$

We can therefore write (dropping in and out labels since we are now calculating perturbatively)

$$\langle\nu, A'|\hat{\mathcal{J}}_\mu|\mu, A\rangle = \frac{G_F}{\sqrt{2}}\int d^4 y[-e\langle\nu|T\{\bar{\nu}(y)\gamma^\nu(1-\gamma_5)\ell(y)\bar{\ell}(0)\gamma^\nu\ell(0)\}|\mu\rangle\langle A'|\hat{J}_\nu(y)|A\rangle$$

$$+ \langle\nu|\bar{\nu}(y)\gamma^\nu(1-\gamma_5)\ell(y)|\mu\rangle\langle A'|T\{\hat{J}_\nu(y)\hat{\mathcal{J}}_\mu(0)\}|A\rangle], \quad (B5)$$

where we have used Eq. (B4) and the fact that $\langle A'|T\{\hat{J}_\nu(y)\hat{\mathcal{J}}_\mu(0)\}|A\rangle = \langle A'|T\{\hat{J}_\nu(y)\hat{\mathcal{H}}_\mu(0)\}|A\rangle$.

The first term can be reduced to a single weak-current nuclear matrix element (measurable with e.g., neutrino scattering) and an electroweak Feynman diagram with the muon radiating a photon. In the second term the nucleus radiates a photon and is excited by a weak current insertion. In between these two current insertions a virtual excitation of some low-lying nuclear energy level could, naïvely, supply an energetic denominator that is $\Delta E \sim O(\text{few MeV})$. Our focus therefore shifts to this term.

Translating the lepton fields to $t = 0$ and $y = 0$ and introducing $Q_\ell = k - k'$ (the four momentum transferred out of the leptons) we see that all of the nuclear physics that could enhance terms proportional to m_*^2 is buried in the current-current correlation function evaluated between two nuclear states

$$W_{\mu\nu} = \int d^4 y \langle A'|T\{\hat{J}_\nu(y)\hat{\mathcal{J}}_\mu(0)\}|A\rangle e^{-iQ_\ell y}. \quad (B6)$$

In terms of this object the second term in Eq. (B5) can be expressed as $\frac{1}{\sqrt{2}}G_F\bar{u}\gamma^\nu(1-\gamma_5)uW_{\mu\nu}$. Let us consider each time ordering separately:

$$\mathcal{W}_{\mu\nu}^+ = \int_0^\infty dy_0 \int d^3 y \langle A'|\hat{J}_\nu(y)\hat{\mathcal{J}}_\mu(0)|A\rangle e^{-iQ_\ell y}, \quad (B7)$$

$$\mathcal{W}_{\mu\nu}^- = \int_{-\infty}^0 dy_0 \int d^3 y \langle A'|\hat{\mathcal{J}}_\mu(0)\hat{J}_\nu(y)|A\rangle e^{-iQ_\ell y}. \quad (B8)$$

Our approach is to
  (i) Insert a complete set of energy eigenstates $\sum_X |X\rangle\langle X| = 1$; we will refer to these states as isobars.
  (ii) Integrate over space to get a momentum conserving delta function.

  (iii) Integrate over time to get denominators of $1/(E \pm i\eta)$.
We find

$$\mathcal{W}_{\mu\nu}^+ = \sum_X \int (2\pi)^3 \delta^{(3)}(\vec{Q}_\ell + \vec{P}_X - \vec{p}')\frac{\langle A'|\hat{J}_\nu|X\rangle\langle X|\hat{\mathcal{J}}_\mu|A\rangle}{(Q_\ell + P_X - p')_0 - i\eta}, \quad (B9)$$

$$\mathcal{W}_{\mu\nu}^- = -\sum_X \int (2\pi)^3 \delta^{(3)}(\vec{Q}_\ell - \vec{P}_X + \vec{p})\frac{\langle A'|\hat{\mathcal{J}}_\mu|X\rangle\langle X|\hat{J}_\nu|A\rangle}{(Q_\ell + p - P_X)_0 + i\eta}. \quad (B10)$$

Next, we can make use of Lorentz invariance to rewrite $\sum_X = \sum_n \int \widetilde{dp}_n$ where states are now labeled by their quantum number n (which determines the mass and spin of the isobar state) and their three momentum $|X\rangle = |n, \vec{P}\rangle$; here $\widetilde{dp} = \frac{1}{2E_p}d^3 p/(2\pi)^3$. Let us introduce $|Z\rangle = |n, \vec{P} = \vec{p}' - \vec{Q}_\ell\rangle$ and $|Y\rangle = |n, \vec{P} = \vec{Q}_\ell - \vec{p}\rangle$ such that

$$W_{\mu\nu} = \sum_Z \frac{\langle A'|\hat{J}_\nu|Z\rangle\langle Z|\hat{\mathcal{J}}_\mu|A\rangle}{2E_Z(\Delta E_\ell + [E_Z - E_{A'}]) - i\eta}$$

$$- \sum_Y \frac{\langle A'|\hat{\mathcal{J}}_\mu|Y\rangle\langle Y|\hat{J}_\nu|A\rangle}{2E_Y(\Delta E_\ell - [E_Y - E_A]) + i\eta}. \quad (B11)$$

We have introduced $\Delta E_\ell = (Q_\ell)_0$ and the energies of the isobars E_Y and E_Z. Notice that the isobars $|Z\rangle$ and $|Y\rangle$ are discrete states whose momentum is fixed by three-momentum conservation, $\vec{P}_Z = \vec{p}' - \vec{Q}_\ell$ and $\vec{P}_Y = \vec{Q}_\ell - \vec{p}$. Small nuclear energy level splittings from the ground state do not lead to small denominators in Eq. (B11). $\Delta E_\ell \sim O(q_0) \sim 100$ MeV is sufficiently large that any low-lying nuclear states are far off shell. The only small denominators that can

appear are in the second time ordering when $[E_Y - E_A] \approx \Delta E_\ell$; such considerations are beyond the scope of this work, and we do not consider them here.

APPENDIX C: COMPARISON WITH KROLL AND WADA

The main difference between [23] (KW) and our study is that we focus on the end point *specifically* and carefully consider the validity of approximating virtual photons by real ones. Since the authors of KW simply assume this approximation as an ansatz many of our results are similar to theirs. We disagree with the authors on certain technical points, and in particular on the validity of their advocated approximation in generic regions of phase space. We find, however, that the ansatz proposed in KW becomes a well controlled approximation in the limit that the electron's kinetic energy becomes much smaller than the virtual photon energy, i.e., when $\delta \ll 1$.

The notation of KW differs substantially from modern treatments and specifically this paper. Moreover, they consider a two-body system and so details of the phase space are slightly different between our works. For the benefit of the interested reader we spell out the differences in notation, derive the results of KW, and comment on typos and conceptual differences.

Kroll and Wada consider a system A, initially at rest, that emits a photon becoming system B of mass M. The mass of system A is parametrized by $M_A = M + E_{AB}$ with E_{AB} the energy splitting between A and B. An important identity is that

$$\frac{1}{2E_{B*}} = \frac{E_{AB} + M}{(E_{AB} + M)^2 + M^2 - m_*^2}. \quad \text{(C1)}$$

The case of a massless photon, E_B, is recovered by setting $m_* = 0$.

In Eq. (2) of KW, the rate of photon emission is calculated. To emphasize the difference in definition we denote our equivalent expression by (P&H) and the expressions of KW by (K&W),

$$\Gamma_\gamma = \frac{1}{8\pi^2} \frac{q_0 M_A}{E_B} \int d\Omega (J_0^{11} + J_0^{22}) \quad \text{(K \& W)}, \quad \text{(C2)}$$

$$\Gamma_\gamma = \frac{1}{2M_A} \times \frac{\bar{\beta}(M_A, 0, M_B)}{32\pi^2} \int d\Omega (\mathcal{J}_0^{11} + \mathcal{J}_0^{22}) \quad \text{(P \& H)}, \quad \text{(C3)}$$

where (borrowing notation from [26])

$$\bar{\beta}(s, m_1, m_2) = \sqrt{1 - \frac{2(m_1^2 + m_2^2)}{s} + \frac{(m_1^2 - m_2^2)^2}{s^2}}. \quad \text{(C4)}$$

For $m_1 = 0$ we have that $\bar{\beta} = 2E_1/M_A$ such that $\bar{\beta}(M_A, 0, M_B) = 2q_0/M_A$ and therefore

$$\Gamma_\gamma = \frac{1}{2M_A} \frac{1}{16\pi^2} \frac{q_0}{M_A} \int d\Omega (\mathcal{J}_0^{11} + \mathcal{J}_0^{22}) \quad \text{(P \& H)}. \quad \text{(C5)}$$

This allows us to identify the (K&W) current with our (P&H) matrix element via

$$(\text{K \& W}) \quad J_\mu = \frac{\sqrt{E_B}}{2M_A} \langle \mathcal{J}_\mu \rangle \quad \text{(P \& H)}. \quad \text{(C6)}$$

This means that there are implicit factors of E_B contained in the expressions in KW, with $J_*^{\mu\nu} \propto E_{B*} \times \mathcal{J}_*^{\mu\nu}$.

Having identified the appropriate current we can now jump ahead to the internal conversion coefficient ρ which is defined as the rate of electron-positron pair production to photon emission. KW gives (with expressions adapted to our notation) in Eq. (8)

$$\rho = \frac{\alpha}{2\pi} \int_{2m_e}^{q_0} \frac{dm_*}{m_*} \int_{-1}^{1} d\cos\vartheta \frac{E_{B*}}{E_B} \frac{|\vec{q}_*|}{|\vec{q}|} \beta_e$$
$$\times \left\{ [2 + 2(\cos^2\vartheta - 1)\beta_e^2] R_T(m_*) + 2[1 - \beta_e^2 \cos^2\vartheta] \right.$$
$$\times \left. \frac{m_*^2}{q_0^2} R_L(m_*) \right\} \quad \text{(K \& W)}, \quad \text{(C7)}$$

where we have used

$$q_0 = \frac{E_{AB}^2 + 2E_{AB}M + m_*^2}{2(E_{AB} + M)}. \quad \text{(C8)}$$

Modifying our discussion in the main text to account for two-body phase space we find that

$$\Phi_{2*} = \Phi_2 \times \frac{\bar{\beta}_*}{\bar{\beta}_0}, \quad \text{(C9)}$$

with

$$\bar{\beta}_* = \sqrt{1 - \frac{2(M_B^2 + m_*^2)}{M_A^2} + \frac{(M_B^2 - m_*^2)^2}{M_A^4}}. \quad \text{(C10)}$$

One can check that $|\vec{q}_*|/|\vec{q}_0| = \bar{\beta}_*/\bar{\beta}_0$. We therefore find [using $dm_*^2 = 2m_* dm_*$ and keeping in mind that we are *not* including the delta function $\delta(\mathfrak{E}_+ - E_+)$],

$$\rho = \frac{\alpha}{2\pi} \int_{2m_e}^{q_0} \frac{dm_*}{m_*} \int_{-1}^{1} d\cos\vartheta$$
$$\frac{|\vec{q}_*|}{|\vec{q}|} \times \beta_e \{ [2 + 2(\cos^2\vartheta - 1)\beta_e^2](1 + T_*^2)$$
$$+ 2[1 - \beta_e^2 \cos^2\vartheta] L_*^2 \} \quad \text{(P \& H)}. \quad \text{(C11)}$$

The factor of E_{B*}/E_B obtained in KW is absent from our expression because our functions T_*^2 and L_*^2 are defined in

terms of $\mathcal{J}$ whereas R_T and R_L are defined in terms of KW's J. As emphasized in Eq. (C6), upon being squared these differ by a factor of E_B, and hence in KW's expression there *should* be a ratio of E_B/E_{B*} instead of the factor of E_{B*}/E_B that appears in Eqs. (8) and (9) of KW [explicitly, $[(E+M)^2 + M^2 - x^2]/[(E+M)^2 + M^2]$ in the notation of KW]; we ascribe this to a typo in KW.

Despite appearances there is no discrepancy between our longitudinal matrix element and those in KW, because KW have evaluated their matrix elements in the lab frame, whereas our expression is evaluated in the rest frame of the virtual photon (or equivalently the electron positron pair). This accounts for the factor of m_*^2/q_0^2 in KW, which can be interpreted as arising from boosting the rest-frame expression into the lab frame.

This final point is important since the naïve m_*^2 suppression appearing in front of $R_L(m_*)$ is artificial, but was used as a justification for the neglect of longitudinal matrix elements in KW. The fact that this is an artificial suppression can be seen in a simple example such as $\ell^+\ell^- \to e^+e^-$ where we can interpret the $\ell^+\ell^-$ pair as a current sourcing a virtual photon γ_*. In this case a straightforward calculation shows that $\mathcal{J}_*^{\mu\nu} = (k^\mu p^\nu + p^\mu k^\nu - \frac{1}{2}sg_{\mu\nu})$ such that $\mathcal{J}_*^{33} = 2m_\ell^2$ in the rest frame (i.e., center of mass frame), whereas it equals $\mathcal{J}_*^{33} = 2\gamma_*^2 m_\ell^2$ in the lab frame, where $\gamma_* = q_0^2/m_*^2$ is the boost of the virtual photon. Note, however, that in *any frame* the combination $m_*^2/q_0^2 \times \mathcal{J}_*^{33} = 2m_\ell^2$, which does *not* vanish in the $m_* \to 0$ limit. This illustrates that longitudinal matrix elements do not decouple in the $m_* \to 0$ limit. Instead, as we have shown in Eqs. (27), (31), and (35), they decouple when $\beta_e \cos\theta \to 1$.

The neglect of finite-m_*^2 corrections to the transverse matrix elements is only valid if $m_*^2 \ll q_0^2$. The integration measure is logarithmic in m_*, and so all scales contribute equally. Therefore, unless $(m_*^{(+)}/q_0)^2 \ll 1$, there is no reason to expect a real-photon approximation to be accurate. Additionally, as we have emphasized in the paragraph above, there is no small-virtuality suppression of longitudinal matrix elements. Rather, these are suppressed when $[1 - \beta_e^2 \cos^2\vartheta]$ is small. As the positron energy approaches its end point, both of these conditions are satisfied with errors being $O(\delta)$.

APPENDIX D: GENERALIZATION TO ARBITRARY FINAL STATES

In Eq. (2) we have assumed a definite exclusive final state. As mentioned in the first footnote of the paper this assumption can be easily relaxed to accommodate inclusive final states for reactions of the form

$$|A\mu, \text{atom}\rangle \to \sum_X |\nu, \gamma, X\rangle. \tag{D1}$$

The expressions for the real photon spectrum and internal e^+e^- spectrum are given by

$$\Gamma_\gamma = \frac{1}{2M_{\text{atom}}} \sum_X \int d\Phi_3(X) \mathcal{J}_0^{\mu\nu}(X)(-g_{\mu\nu}^\perp), \tag{D2}$$

$$\Gamma_{e_+e_-} = \frac{1}{2M_{\text{atom}}} \sum_X \int d\Phi_4(X) \mathcal{J}_*^{\mu\nu}(X) \frac{4\pi\alpha}{m_*^4} L_{\mu\nu}, \tag{D3}$$

where we have included an X dependent evaluation of the electromagnetic current. Four-body phase space can be decomposed for each final state X as described in the main text. If X is has a multiparticle final state then an analogous decomposition for n-body phase space can be performed.

Working at leading order in δ the resultant e^+e^- spectrum can then be related to the photon spectrum by the same internal conversion probability for every final state X

$$\frac{d\Gamma_{e_+e_-}}{dE_+} = \int dE_\gamma \sum_X \frac{d\Gamma_X}{dE_\gamma} P_{\text{int}}^{(0)}(E_+|E_\gamma). \tag{D4}$$

The full differential photon flux is then easily identified as $d\Gamma = \sum_X d\Gamma_X$, such that

$$\frac{d\Gamma_{e_+e_-}}{dE_+} = \int dE_\gamma \frac{d\Gamma}{dE_\gamma} P_{\text{int}}^{(0)}(E_+|E_\gamma). \tag{D5}$$

This agrees with Eq. (36).

[1] Y. Kuno and Y. Okada, Muon decay and physics beyond the standard model, Rev. Mod. Phys. **73**, 151 (2001).

[2] W. J. Marciano, T. Mori, and J. M. Roney, Charged lepton flavor violation experiments, Annu. Rev. Nucl. Part. Sci. **58**, 315 (2008).

[3] R. H. Bernstein and P. S. Cooper, Charged lepton flavor violation: An experimenter's guide, Phys. Rep. **532**, 27 (2013).

[4] A. de Gouvea and P. Vogel, Lepton flavor and number conservation, and physics beyond the standard model, Prog. Part. Nucl. Phys. **71**, 75 (2013).

[5] L. Calibbi and G. Signorelli, Charged lepton flavour violation: An experimental and theoretical introduction, Riv. Nuovo Cimento **41,** 71 (2018).

[6] L. Galli, Lepton Flavour violation in muon decays, in *Proceedings of the 17th Conference on Flavor Physics and CP Violation* (2019) [arXiv:1906.10483].

[7] J. Kaulard *et al.* (SINDRUM II Collaboration), Improved limit on the branching ratio of $\mu^- \to e^+$ conversion on titanium, Phys. Lett. B **422,** 334 (1998).

[8] W. H. Bertl *et al.* (SINDRUM II Collaboration), A Search for muon to electron conversion in muonic gold, Eur. Phys. J. C **47,** 337 (2006).

[9] R. H. Bernstein (Mu2e Collaboration), The Mu2e experiment, Front. Phys. **7,** 1 (2019).

[10] M. J. Lee, COMET muon conversion experiment in J-PARC, Front. Phys. **6,** 1 (2018).

[11] T. Geib, A. Merle, and K. Zuber, $\mu^- - e^+$ conversion in upcoming LFV experiments, Phys. Lett. B **764,** 157 (2017).

[12] A. De Gouvêa, W.-C. Huang, J. König, and M. Sen, Accessible Lepton-Number-Violating models and negligible neutrino masses, Phys. Rev. D **100,** 075033 (2019).

[13] J. M. Berryman, A. de Gouvêa, K. J. Kelly, and A. Kobach, Lepton-number-violating searches for muon to positron conversion, Phys. Rev. D **95,** 115010 (2017).

[14] F. T. Avignone III, S. R. Elliott, and J. Engel, Double beta decay, majorana neutrinos, and neutrino mass, Rev. Mod. Phys. **80,** 481 (2008).

[15] B. Yeo, Y. Kuno, M. J. Lee, and K. Zuber, Future experimental improvement for the search of lepton-number-violating processes in the $e\mu$ sector, Phys. Rev. D **96,** 075027 (2017).

[16] T. Meissner, F. Myhrer, and K. Kubodera, Radiative muon capture by a proton in chiral perturbation theory, Phys. Lett. B **416,** 36 (1998).

[17] S.-I. Ando and D.-P. Min, Radiative muon capture in heavy baryon chiral perturbation theory, Phys. Lett. B **417,** 177 (1998).

[18] H. W. Fearing and G. E. Walker, Radiative muon capture in a relativistic mean field theory: Fermi gas model, Phys. Rev. C **39,** 2349 (1989).

[19] H. W. Fearing and M. S. Welsh, Radiative muon capture in medium heavy nuclei in a relativistic mean field theory model, Phys. Rev. C **46,** 2077 (1992).

[20] P. C. Bergbusch *et al.*, Radiative muon capture on O, Al, Si, Ti, Zr, and Ag, Phys. Rev. C **59,** 2853 (1999).

[21] M. K. Cheoun, K. S. Kim, and T. K. Choi, Radiative muon capture and induced pseudoscalar coupling constant in nuclear matter, J. Phys. G **29,** 2099 (2003).

[22] P. Christillin, M. Rosa-Clot, and S. Servadio, Radiative muon capture in medium heavy nuclei, Nucl. Phys. **A345,** 331 (1980).

[23] N. M. Kroll and W. Wada, Internal pair production associated with the emission of high-energy gamma rays, Phys. Rev. **98,** 1355 (1955).

[24] H. W. Fearing, A relativistic calculation of radiative muon capture in hydrogen and ^{3}He, Phys. Rev. C **21,** 1951 (1980).

[25] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, USA, 1995).

[26] H. Murayama, Notes on phase space, http://hitoshi.berkeley.edu/233B/phasespace.pdf, accessed: 2020-05-29.

[27] D. H. Wilkinson, Analysis OF neutron beta decay, Nucl. Phys. **A377,** 474 (1982).

[28] P. Schlüter, G. Soff, and W. Greiner, Pair creation by internal conversion, Phys. Rep. **75,** 327 (1981).

[29] R. Plestid and R. J. Hill (to be published).