We thank you for your very careful reading of our article. We list here our answers to your questions: We hope this will clear some of the issues. We will be pleased to reply to any further question you may have.

Q1: *You seem to use the "soft wall" model in which $\phi(z) = \lambda z^2$ [28]. We could not find what the justification was for this choice.*

A1: In older publications the use of a "soft wall" dilaton $\phi(z) = \lambda z^2$ in the meson AdS Lagrangian (and a Yukawa-like term for baryons) was indeed an *ad hoc* choice, only justified by its phenomenological success, namely to generate linear Regge trajectories for mesons and baryons. The constraints imposed by the superconformal algebraic structure, as decribed in detail in [26] and [27] fix, however, the form of the interaction and lead precisely to the quadratic dilaton $\phi(z) = \lambda z^2$ for bosonic states, and a Yukawa interaction for baryon states (see also point 2 below). Furthermore, this procedure leads to a precise relation between the meson and baryon spectrum, as described in [27].

It is interesting to notice that the procedure introduced by de Alfaro, Furlan and Fubini, which was generalized to superconformal algebra by Fubini and Rabinovici, shows how a mass scale can appear in the Hamiltonian while maintaining the conformal invariance of the action. This procedure also leads to a massless pion in the $m_q \to 0$ chiral limit.

Q2: *Our understanding of the conformal theory is that it means that the coupling doesn't run with scale. In that sense, how does one justify applying DGLAP evolution on top of the conformal result? It sounds like the more consistent thing to do would be to compute the running with the scale from the theory directly as corrections to the conformal limit.*

A2: This is an interesting proposal for a truly conformal theory. However, in the LF holographic QCD framework (LFHQCD), conformal quantum mechanics is used to introduce a mass scale. This was subsequently extended to superconformal quantum mechanics, which provides a precise form of the confinement potential [26, 27] and thus determines the dilaton profile. Therefore, the result is a non-conformal model with a mass scale and confinement. We use this structural framework, to compute the GPDs at the hadron scale. These are the initial quark distributions which are then evolved to higher scales using the usual procedures.

Q3: *The quantity e_5 is chosen to be $e_5(z) = \exp[(\lambda + |\lambda|)z^2/2]e$, however, we could not find a derivation of this result.*

A3: This z-dependence of the EM charge in the 5-dimensional AdS space is determined by the requirement of charge quantization in the 4-dimensional space. It is explained in detail in [28], Sect. 6.1.4.

Q4: *In the derivation of Eq. (1) for the form factor in terms of the Beta function, the derived result appears to involve $B[\tau - 1, 1 - t/(4\lambda)]$. At some point the second argument gets changed to $(1/2) - t/(4\lambda)$, as in your Eq. (1). We could not find a theoretical justification for this.*

A4: The results from LFHQCD were derived for the specific case of zero AdS mass. From the holographic embedding the AdS mass is related to kinematical quantities by [28] $(\mu R)^2 = -(2-J)^2 + L^2$, thus for $J = 1$ it follows that $L = 1$. From the LFHQCD spectral formula [28]

$$M^2 = 4\lambda \left(n + \frac{J+L}{2} \right),$$

we can deduce the corresponding Regge trajectory for $J = L$, $n = 0$: It is $\alpha(t) = \frac{t}{4\lambda}$, which leads to the result given above for the Beta function.

On the other hand, the ρ meson is a $J = 1, L = 0$ state. Its physical Regge trajectory for $J = L+1$, $n = 0$, also follows from the spectral formula above and it is given by $\alpha(t) = \frac{1}{2} + \frac{t}{4\lambda}$, the result in Eq. (1). This pole shift in the amplitudes described in Refs. [28] and [29] does not modify the general expression for the form factor $F(t) \sim B(\tau - 1, 1 - \alpha(t))$, which remarkably has identical Veneziano-type structure [59].

Q5: *The matching scale is chosen to be $\mu_0 = 1.06 \pm 0.15$ GeV, with reference to [75]. However, the scale there is 1.14 ± 0.12 GeV. Was the new value determined from the present analysis?*

A5: In fact no: The value of μ_0 is independent of the present analysis. The value of μ_0 differs from the value quoted in Ref. [75] because it depends on the order of the pQCD series which is considered. The value in Ref. [75] is determined at N^{4}LO while the present work is done at NNLO. When the matching procedure of [75] is performed at NNLO, one obtains $\mu_0 = 1.06$ GeV. Please note that although 1.14 ± 0.12 GeV and 1.06 ± 0.12 GeV are compatible, we conservatively increased the uncertainty to account for the order dependence.

Q6: *Can you please explain how to split Eq. (8) into the PDF and "profile function" and give a physical interpretation for the latter?*

A6: We follow the standard notation (See for example [66]), but here the separation into the PDF $q(x)$ and the profile function $f(x)$ comes out naturally and are both expressed in terms of the function $w(x)$. The function $f(x)$ describes the change of the profile of the quark distribution in transverse impact space, $a_\perp = (1-x)b_\perp$, with respect to x, as given by the effective LF wave function (22).

Q7: *From the discussion on p. 2 of your paper, it looks like the function $w(x)$ is constructed in order for the PDF to have specific behaviors in the $x \to 0$ and $x \to 1$ limits. In particular, the behavior $w(x) \sim x$ is chosen to give you the Regge behavior for the valence PDF, $q(x) \sim x^{-1/2}$. Is there any derivation of the $w(x) \sim x$ behavior in your model, or is it just chosen in order to reproduce the Regge result?*

A7: The reparametrization function $w(x)$ obeys the constraints (7). Therefore, one starts with the general ansatz $w(x) = C\, x^\beta$ where $\beta > 0$. The only possible Regge trajectories are the ρ and the (soft) pomeron trajectory. No choice of β can lead to the Pomeron trajectory, whereas $\beta = 1$ is fully compatible with the degenerate ρ trajectory.

Q8: *Is the function $w(x)$ a universal function, or would it have to be different, say, for sea quarks, which would have a behavior closer to $q(x) \sim x^{-1}$, requiring $w(x) \sim x^2$?*

A: We have not explored yet the sea quark distribution in the present framework, but from our experience with the valence distributions it is clear that the $x \to 0$ behavior depends critically on the Regge intercept, $x^{-\alpha(0)}$, which determines the $x^{-1/2}$ dependence for the quark valence distribution. Therefore a behavior $q(x) \sim x^{-1}$ would rather hint at an intercept $\alpha(0) \sim 1$ –characteristic of the Pomeron, while maintaining universality in $w(x)$. It is worth examining this issue further and we thank you for the comment.

Q9: *For the $x \to 1$ behavior, $w(x)$ is chosen such that $1 - w(x) \sim (1-x)^2$. This gives you $q(x) \sim (1-x)$ for the pion and $\sim (1-x)^3$ for the nucleon. According to the counting rule arguments, for the pion you should have an additional power of $(1-x)^{2*\Delta s_z}$, where Δs_z is the difference between the helicities of the parton and hadron. In the case of the pion, it gives an additional power of $(1-x)$. Is this not present in your holographic model?*

A9: Right, we impose the condition $w'(1) = 0$ which leads to the nonperturbative result $q(x) \sim (1-x)^{2\tau-3}$, the Drell-Yan local counting rule [61] for the leading-twist τ term. The additional power of $1 - x$ from opposite quark and hadron helicities in perturbative QCD is not present in our nonperturbative framework.

Q10: *What does your model predict for the x dependence of the sea quark and gluon PDFs?*

A10: We have not explored this problem yet.

Q11: *The use of the term "twist" for the parameter τ is confusing. The leading term in the nucleon PDF, for instance, has $\tau = 3$, whereas in the usual definition of "twist" in operator language both the pion and nucleon PDFs are of twist 2, so τ cannot correspond to what is usually known as "twist".*

A11: Right: Twist τ in the article differs from the conventional notion of "twist" for parton distributions (the valence distributions discussed in the article are all twist 2). Twist τ here is the canonical dimension minus spin, $\tau = \Delta - \sigma$, where σ is the sum over the constituents spin, $\sigma = \sum_{i=1}^{n} \sigma_i$: It is also equal to the number of constituents in a given Fock component in the Fock expansion of a hadron state. It thus controls the short distance behavior of the hadronic state and thus the power-law asymptotic behavior of form factors and amplitudes as first shown in Refs. [53] and [54].

Q12: *The nucleon PDFs in (16)-(17) are constructed from a $\tau = 3$ term and a $\tau = 4$ term (which gives a $(1-x)^5$ behavior for the d_v PDF in Eq. (17) with your choice of $r = 3/2$). Is the choice $r = 3/2$ made in order to get a faster fall-off at high x of the d_v PDF than the $\tau = 3$ term would give you?*

A12: Indeed no: in [56] we found that the value of r should be around 2. However when we attempted to describe the PDFs using the present framework we found an additional constraint on the maximum possible value of r: It should be 3/2, otherwise d_v becomes negative. Thus we took the maximum possible value for r to have the best possible description of the form factors. It turns out that taking this value the twist 3 contribution to the d_v distribution vanishes.

Q13: *In Eq. (20), you have a $\tau = 2$ term and a $\tau = 4$ term. Why is there no $\tau = 3$ term here?*

A13: This comes from the holographic description of the time-like pion form factor given in Refs. [28] and [81]. The twist-3 contribution was found to be very small and dropped out. This is consistent with the fact that constituent gluons are not favored in LFHQCD.

Q14: *With the choice of 2 shapes in (16), (17) and (20) with their respective parameters, it seems to us that what is ultimately done is a phenomenological fit to the data, motivated by some theoretical considerations, rather than a quantitative calculation with predictive power. In this sense we do not understand your claim that this is a "parameter free prediction" in any usual sense of the term.*

A14: The expressions for the distributions in the nucleon, based on (16) and (17) are indeed a very economical (single-free parameter) fit, based on the important structural property (8) and theoretical constraints (Regge behavior and local counting rules) near the end points $x = 0$ and 1. Since the x dependence is fixed by this equation, the application to the distribution in the pion, however, needs no additional parameter: It is a prediction.