1 Convective Penetration argument

First some definitions:

1. $z = L_s$ is the the top of the convection zone according to the Schwarzschild criterion; it's the height where $\boldsymbol{\nabla}_{\text{ad}} = \boldsymbol{\nabla}_{\text{rad}}$.
2. L_{cz} is the top of the convection zone; it's roughly the top of the region where convection flattens $\boldsymbol{\nabla} \to \boldsymbol{\nabla}_{\text{ad}}$. We generally get $L_{\text{cz}} > L_s$.
3. The penetration depth is $\delta_p = L_{\text{cz}} - L_s$.
4. The flux carried by convection for $z < L_s$ is $F_{\text{conv,cz}} = Q\delta_H$, where Q is the magnitude of the internal heating and δ_H is the depth of the heating layer.
5. $\overline{w}$ is the vertical profile of the characteristic (vertical) convective velocity, which is a constant w_{cz} for $z \leq L_{\text{cz}}$.
6. Similarly, $\overline{\delta T}$ is the vertical profile of the characteristic temperature perturbation.

And some key assumptions.

1. Convection flattens $\boldsymbol{\nabla} \to \boldsymbol{\nabla}_{\text{ad}}$ for $z \leq L_{\text{cz}}$ (baked into our definitions).
2. We assume a system in thermal equilibrium, at least in (adiabatic) convection zone with $z \leq L_{\text{cz}}$. Therefore $F_{\text{conv}}(z) = F_{\text{tot}}(z) - F_{\text{rad,ad}}(z)$, where F is a flux and $F_{\text{rad,ad}}$ is the radiative flux along the adiabatic gradient.
3. We assume $F_{\text{conv}} = \overline{wT} \approx \overline{w}\,\overline{\delta T}$. Combined with our previous assumption, we get
$$\overline{\delta T} \approx \frac{F_{\text{tot}}(z) - F_{\text{rad,ad}}(z)}{\overline{w}}. \tag{1}$$

As we increase in height, the conductivity and thus $F_{\text{rad,ad}}(z)$ also increases. This means that $\overline{\delta T}$ has the opposite sign of $\overline{w}$ for $z > L_s$. We presume that buoyancy breaking is the dominant mechanism which brings convective motions to a stop in this adiabatic layer,

$$\frac{d\overline{w}}{dt} = \overline{\delta T} = \frac{F_{\text{tot}}(z) - F_{\text{rad,ad}}(z)}{\overline{w}}. \tag{2}$$

If we multiply both sides by $\overline{w}$ and absorb it into the the derivative, then apply the chain rule with $d/dt = d/dz(dz/dt) = d/dz\overline{w}$, we retrieve,

$$\frac{1}{2}d\overline{w}^3 = \left[F_{\text{tot}}(z) - F_{\text{rad,ad}}(z)\right]dz. \tag{3}$$

We now replace $F_{\text{rad,ad}}(z) = k_0(z)\boldsymbol{\nabla}_{\text{ad}}$, and we assume that $k(z)$ instantaneously jumps from a low value k_{cz} to a high value k_{rz} at $z = L_s$. Per this assumption, the convective flux is a negative constant for $z \geq L_s$. We integrate from $z = L_s$ with $\overline{w} = w_{\text{cz}}$ to $z = L_s$ with $\overline{w} = 0$, and get

$$-\frac{w_{\text{cz}}^3}{2} = F_{\text{conv,p}}\delta_p. \tag{4}$$

By definition, the convective flux in the penetrative layer with $z > L_s$ is related to $\mathcal{P}$,

$$F_{\text{conv,t}} = -\frac{F_{\text{conv,cz}}}{\mathcal{P}}, \tag{5}$$

so we retrieve

$$\delta_p = \frac{w_{\text{cz}}^3\mathcal{P}}{2F_{\text{conv,cz}}}, \tag{6}$$

and the penetration depth scales with the cube of the velocity and linearly with $\mathcal{P}$. This prediction does not exactly line up with the simulation results (this overpredicts the magnitude of δ_p). The important thing we need to determine is if it scales with w_{cz}, $\mathcal{P}$, and $F_{\text{conv,cz}}$ in the appropriate way. So far, it seems to scale linearly with $\mathcal{P}$ at fixed other parameters.