Revista Mexicana de Astronomía y Astrofísica, **45**, 107–123 (2009)

ON THE ABUNDANCE ENIGMA IN IONIZED REGIONS

J. Bohigas

Instituto de Astronomía
Universidad Nacional Autónoma de México, Ensenada, B. C., Mexico

Received 2009 January 20; accepted 2009 February 20

RESUMEN

En regiones ionizadas con gradientes y fluctuaciones de temperatura el cociente de la abundancia iónica determinada con una línea de recombinación entre la abundancia iónica encontrada con una línea de excitación colisional (CEL), o ADF, es menor a lo observado (ADF ≥ 2). Mayores ADFs son posibles si hay otra componente $\geq 30\%$ más fría. La temperatura en la región fría es ≈ 500, 200 y 100 K si el ADF encontrado de un línea IR $\simeq 2$, 5 y 10. La mayor parte de la masa está en la región caliente. La masa total de H^+ se subestimó si fue determinada de la intensidad de una línea Balmer. Las imágenes de [O IIII]5007/Hβ también muestran la distribución relativa de masa caliente y fría. La medición de abundancias se dificulta por la imposibilidad de separar la luz de las componentes fría y caliente, por la existencia de diversos conjuntos de abundancias y por la insuficiente información espectral de la región caliente.

ABSTRACT

In ionized regions with temperature gradients and fluctuations, the ratio of the ion abundance obtained from a recombination line to that found from a collisionally excited line (CEL), or ADF, is smaller than observed (ADF ≥ 2). Larger ADFs are found when there is an additional component that is $\geq 30\%$ colder. The temperature in the cold component must be ≈ 500, 200 and 100 K if the ADF found from an IR CEL is $\simeq 2$, 5 and 10. Most of the mass is in the hot region. The total H^+ mass has been underestimated if it was found from the intensity of a Balmer line. [O IIII]5007/Hβ images can also render the relative distribution of cold and hot matter. The determination of accurate abundances is forestalled by the fact that observations cannot discriminate light from these components, the existence of distinct abundance sets and insufficient spectral information for the hot region.

Key Words: ISM: abundances — ISM: H II regions — ISM: planetary nebulae, general

1. INTRODUCTION

In emission nebulae, ion abundances of heavy elements relative to H^+ have been usually determined from the intensity of a collisionally excited line (henceforth CEL) relative to the intensity of a Balmer series line. In the standard approach, both line emissivities are calculated assuming that temperature –as well as density, ion concentration, abundance and other physical variables– is uniform in the emission region. Since UV and optical CELs have large excitation energies and their emissivity has an exponential temperature dependence, a minor inaccuracy or variation in the temperature will spread out as a large uncertainty in the ion abundance. This is not an issue when a recombination line (henceforth RL) is used, since in this case the emissivity has a power law temperature dependence which almost cancels out with the power law term for the Balmer line emissivity, and temperature uncertainties have a smaller effect on the accuracy with which ion abundances are determined. The problem with RLs is that these can be hundreds of times weaker than CELs and the spectral resolution is not always large enough to separate individual lines. Thus, abundances found from RLs can be inaccurate due to insufficient spectral resolution and signal-to-noise ratio. This limitation was overcome with the advent of high resolution spectrographs attached to large telescopes using sensitive detectors.

Heavy element ion abundances from RLs have also been determined applying the standard approach. It has been found that the ratio of the ion abundance found from RLs to the ion abundance found from CELs (hereon the abundance discrepancy factor or ADF) is always larger than one. For instance, in H II regions such as Orion, M8 and M17, the O^{+2} abundance found from O^{+} RLs is about two times larger than that found from O^{+2} CELs, assuming the same temperature in both cases (Peimbert, Storey, & Torres-Peimbert 1993; Esteban et al. 1998; Esteban et al. 1999; García-Rojas et al. 2007). The same effect has been observed for an equally long time in planetary nebulae (Kholtygin & Feklistova 1992; Peimbert et al. 1993; Nikitin et al. 1994), where the CEL *vs.* RL abundance discrepancy has been studied in many more objects (see reference lists in Liu et al. 2004, Tsamis et al. 2004 and Robertson-Tessi & Garnett 2005). As with H II regions, the ADF for O^{+2} is close to 2 most of the time, but ADF > 5 in $\sim 20\%$ of those PNs where this quantity has been measured. The fact that ADFs in H II regions and most PNs are so similar, indicates that this effect is primarily related to some unaccounted factor in the way emissivites are calculated or ionized nebulae modeled. It must be pointed out that all of the above has also been found for other ions besides O^{+2}.

The underlying cause for the existence of the ADF is still debated. Some of the lines of research that have been pursued are:

(1) Besides recombination, other mechanisms can produce permitted lines, such as line and continuum fluorescence or radiative charge transfer, and the excitation mechanism has to be identified for each line and circumstance. In order to establish if recombination is the excitation mechanism, observed relative line intensities are compared with theoretical predictions. For example, Grandi (1976) and Esteban et al. (1998, 1999) found that a large number of optical permitted lines were excited by continuum fluorescence, but that all O II lines were produced by recombination, and their intensities implied similar O^{+2} abundances. Recombination lines in PNs are identified in the same manner (e.g., Liu et al. 1995).

(2) ADFs may be an indication that there are large errors in the recombination emissivities. Consistent errors in the radiative recombination coefficients would produce similar ADFs in all emission nebulae. Since this is not observed, ADFs can not be attributed to defective atomic physics (Liu et al. 2000; Robertson-Tessi & Garnett 2005), though there is room for improvement in this research topic (e.g., Zatsarinny et al. 2004).

(3) Four decades ago Peimbert (1967) showed that in the presence of temperature fluctuations about a mean value in an otherwise uniform nebula, most CEL emission will come from the hotter gas, whereas RL emission will be produced in the colder regions. He concluded that if these conditions occur in real nebulae, the standard method applied to CELs will underestimate abundances. This prediction was vindicated with the discovery of ADFs, but the interpretation is not fully convincing. For one thing, the required temperature fluctuations can be quite large: the t^2 parameter must be ≥ 0.04 to account for ADF ≥ 2, which implies rms temperature variations larger than $\sim 20\%$. These have not been identified (see Mathis 1995 and Garnett & Dinnerstein 2001a discussion). But the strongest argument against this scheme is that large ADFs have also been found from FIR CELs (e.g., Liu et al. 2000, 2004), where the effect of temperature fluctuations about $\sim 10^4$ K should be minimal since the excitation energy of these lines is much smaller.

(4) Liu et al. (2000) suggested that most RL emission is produced in a cold hydrogen-poor component immersed in the extended nebula where the bulk of CEL emission is produced. A cold ionized component has been detected in several PNs, where temperatures as low as ~ 500 K have been found from the RL intensity ratio O II 4089/4649 (Tsamis et al. 2004; Liu et al. 2006). This component is visualized as a multitude of blobs or speckles, a concept supported by Tsamis et al. (2008), who found maxima of the ADF at a ~ 1 arcsec scale (less than about 1000 AU) in 3 PNs. On the other hand, Garnett & Dinnerstein (2001b) point out that an inverse correlation between ADF and PN surface brightness would imply that cold speckles condense at late stages. This is not expected, since hydrogen-poor blobs can only be ejected in the final phase of the AGB star and will be subsequently evaporated in the surrounding hot medium. In addition, this component must be He-rich, since PN progenitors cannot enhance C, N, O and Ne simultaneously and by a similar amount (van den Hoek & Groenewegen 1997; Marigo 2001), but so

far there is no evidence for this. In the case of H II regions, the metal-rich component may be made of exogenous inclusions from a nearby supernova event (Tsamis & Pequignot 2005) or from supernova ejecta "droplets" falling back onto the galactic disk (Stasińska et al. 2007). Notice that this scenario is, in some sense, an extreme version of Peimbert's (1967) scheme.

At this point a rigorous definition of the ADF is in order. In the standard approach, abundance, ion concentration and electron density and temperature are assumed to be the same in the regions where ions A^i and H^+ are predominant. This idealized and probably incorrect assumption is used for every ion, but not necessarily using the same temperature for all of them (ionized hydrogen may have many different temperatures, which is obviously inconsistent). Thus, relative ion abundances in the standard approach are given by

$$\frac{N(A^i)}{N(\mathrm{H}^+)} = \frac{I(A^i)E(H, T_0)}{I(\mathrm{H}^+)E(A^i, T_0)} , \tag{1}$$

where $I(\mathrm{H}^+)$ and $I(A^i)$ are line intensities from H^+ and A^i, and $E(\mathrm{H}, T_0)$ and $E(A^i, T_0)$ are the emissivities of these two lines at a temperature T_0. Consequently, the abundance discrepancy factor is given by,

$$\mathrm{ADF} = \frac{I_r(A^i)Q(T_0, A^i)}{I_q(A^i)R(T_0, A^i)} , \tag{2}$$

where I_r and I_q are the intensities of the recombination and collisionally excited lines and R and Q are their emissivities.

This paper started out as an exercise to visualize the consequences that the simplifying assumptions of the standard approach have on the abundance discrepancy, since these are so blatantly limited that it would have been surprising if the ADF had turned to be close to one. A spherically symmetric model for an ionized region with temperature and pressure gradients and fluctuations is described and discussed in § 2. Since no conditions were found where this model can bring forth a mean ADF larger than $\sim$ 1.5–2, the cold component suggested by Liu et al. (2000) is added to the region in § 3, where requirements to produce optical and FIR ADFs between 2 and 20 are also discussed. The possible implications on the way H II regions and PNs are interpreted when a cold component is included are discussed in § 4. Conclusions are given in § 5.

2. GRADIENTS AND FLUCTUATIONS IN AN IONIZED REGION

The line intensity I of a recombination or a collisionally excited line in a plasma occupying a volume V is given by

$$I = \int (N(e)/N)(N(A^i)/N(A))(N(A)/N)N^2E(T)dV , \tag{3}$$

where $N(e)$, $N(A^i)$, $N(A)$ and N are the electron, ion, element and total plasma density, $E(T)$ is the line emissivity and T is the temperature. For an ideal gas this equation can be written as (in units where the Boltzmann constant is one),

$$I = 4\pi \int n(e)n(A^i)n(A)\ (P^2/T^2)\ E(T)r^2dr , \tag{4}$$

where $n(e) = N(e)/N$, $n(A^i) = N(A^i)/N(A)$, $n(A) = N(A)/N$, P is the thermal pressure and r is the distance to the exciting source. Spherical symmetry has been assumed. Obviously, the integration is carried out over the region where $n(A^i) \neq 0$. With $n(e)$, $n(A^i)$, $n(A)$, pressure, temperature and position given in terms of typical or characteristic values for these variables in the ionized region ($n_0(e)$, $n_0(A^i)$, $n_0(A)$, P_0, T_0 and r_0),

$$I = 4\pi r_0^3 n_0(e)n_0(A^i)n_0(A)(P_0/T_0)^2E(T_0) \int ea_ia(p/t)^2E(T)/E(T_0)\xi^2d\xi , \tag{5}$$

where $e = n(e)/n_0(e)$, $a_i = n(A^i)/n_0(A^i)$, $a = n(A)/n_0(A)$, $p = P/P_0$, $t = T/T_0$ and $\xi = r/r_0$. Consequently, if T_0 is the temperature selected in the standard approach, the ADF (Equation 2) is given by

$$\mathrm{ADF} = \frac{\int e\ a_i\ a\ (p^2/t^2)\ R(T)/R(T_0)\xi^2d\xi}{\int e\ a_i\ a\ (p^2/t^2)\ Q(T)/Q(T_0)\xi^2d\xi} , \tag{6}$$

where R and Q are the emissivities of the recombination and collisionally excited lines. It is easy to see that (Osterbrock & Ferland 2006; Storey 1994)

$$\frac{Q(T)}{Q(T_0)} = t^{-0.5} \exp[\chi/T_0(1-1/t)] \,, \tag{7}$$

$$\frac{R(T)}{R(T_0)} = t^b \frac{1+a_1T_0t+a_2T_0^2t^2}{1+a_1T_0+a_2T_0^2} \,, \tag{8}$$

where b, a_1 and a_2 are constants and $\chi = h\nu/k$, ν being the frequency of the collisionally excited line, and h and k the Planck and Boltzmann constants. Notice that Equation (7) holds true only if the electron density is smaller than the critical density; otherwise the emissivity of the collisionally excited line and the ADF are overestimated. This is not an issue with optical CELs, since their critical densities are at least a few times 10^4 cm^{-3}. Not so for all fine structure FIR forbidden lines, where these are closer to 10^3 cm^{-3} at a temperature $\sim 10^4$ K (Osterbrock & Ferland 2006).

The right hand side (hereon RHS) of Equation (6) is equal to one in a nebula where temperature, pressure, density, ion concentration and chemical composition are constant, as assumed in the standard approach. Since the ADF is always larger than one, the situation must be more complex than a single nebula where these quantities are constant.

Suppose that temperature and pressure fluctuations δt and δp, are superimposed on a regular monotonic behavior for these two physical variables, so that $t = t + \delta t$ and $p = p + \delta p$, with $\delta t/t$ and $\delta p/p \ll 1$. The radial distributions of the regular components of temperature and pressure are now assumed to be given by the following power laws:

$$t = \xi^{1/\gamma} \,; \qquad p = \xi^{\beta} \,, \tag{9}$$

where γ can be either positive or negative (temperature may or may not increase away from the exciting star), but β is less or equal zero, since pressure cannot increase away from the exciting star without compromising the dynamical stability of the nebula. These power laws lead to

$$\xi^2 d\xi = \gamma t^{3\gamma-1} dt \,; \qquad p = t^{\gamma\beta} \,, \tag{10}$$

and

$$\mathrm{ADF} = \frac{\int \mathbf{\Gamma}(t+\delta t)^b[1+a_1T_0(t+\delta t)+a_2T_0^2(t+\delta t)^2]/[1+a_1T_0+a_2T_0^2]dt}{\int \mathbf{\Gamma}(t+\delta t)^{-0.5}\exp[\chi(1-1/(t+\delta t))/T_0]dt} \,, \tag{11}$$

where

$$\mathbf{\Gamma} = e \; a_i \; a \; (t^{\gamma\beta}+\delta p)^2 \; t^{3\gamma-1}/(t+\delta t)^2 \,. \tag{12}$$

For each t, temperature fluctuations are supposed to be given by

$$\delta t = \frac{\sum_{k_1}^{k_2} \Omega_k \sin[\Phi_k + (t-0.75)]}{k_2 - k_1} \,, \tag{13}$$

where k_1, k_2, Ω_k and Φ_k are determined randomly, with their values between 1 and 100 (the first two, with $|k_2 - k_1| > 10$), 0.002 and 0.2 (Ω_k) and 0 and $\pi/2$ radians (Φ_k). Fluctuations are normalized so that the largest value of $\delta t/t$ is 0.1. On the other hand, it is assumed that $\delta p = \delta t/m$, with m randomly selected between 1 and 5.

To complete the model, assumptions must be made for the electron, ion and elemental abundance distributions in the region. In the absence of any compelling reason to expect otherwise, the elemental abundance is taken to be constant. The electron density is also assumed constant. This is inadequate for any ion with an ionization potential close to 13.6 eV, since the electron density at the ionization front will change drastically. It is not entirely correct for ions with ionization potentials close to those of He^{+2} and He^{+}, such as C^{+} and O^{+2} (with ionization potentials 24.38 and 54.94 eV respectively), since the electron density can change by roughly 10% in these regions. Finally, the ionized fraction $n(A^i)$ will be close to one except where the ion recombines or is ionized. Numerical photo-ionization models show that this happens over comparatively small distances. The ADF has been calculated assuming that $n(e)$, $n(A^i)$ and $n(A)$ are constant ($e = a_i = a = 1$) and all the following combinations:

(1) Since temperature may increase or decrease away from the exciting star, Equation (11) is computed for $\gamma = -2, -1, 1$ and 2.

(2) Since pressure cannot increase away from the exciting star, Equation (11) is computed for $\beta = -0.5$ and 0.

(3) The relative collisional emissivity (Equation 7) is determined for the collisionally excited lines [O III] 5007 Å and 52 μm.

(4) The relative recombination emissivity (Equation 8) is computed for the $3d - 3p$ transitions of O II recombination lines, so that $b = -0.75885$, $a_1 = -0.22862/10^4$ and $a_2 = 0.02893/10^8$ (Liu et al. 1995).

(5) To cover the observed range of temperatures used in the standard approach, relative emissivities were computed with $T_0 = 9000$, 12000, 15000 and 18000 K.

(6) Integration limits t_1 and t_2 must also be selected. The integration range is kept constant at $t_2 - t_1 = 0.3$ (temperature in the region is assumed to change $\pm 15\%$), but the limits are changed. If T_0 happens to be the mean temperature of the region where $n(A^i) = 1$, then $t_1 = 0.85$ and $t_2 = 1.15$. If T_0 is larger than the the mean temperature, then $t_1 = 0.75$ and $t_2 = 1.05$ is a more adequate selection for the temperature limits. On the other hand, $t_1 = 0.95$ and $t_2 = 1.25$ is the right choice when T_0 is smaller than the mean value. Equation (11) is calculated for these three pairs.

Samples of temperature and pressure distributions, as well as the functional behavior of both integrands, for every combination of power law indices, reference temperature T_0 and integration limits t_1 and t_2, are shown in Figures 1a,b,c,d. The integral for the CEL line was calculated for the [O III] 5007 Å and 52 μ lines. Fluctuations are different in each case since, as mentioned above, the fluctuation parameters are randomly selected.

Numerical integrations, where the CEL is [O III] 5007 Å, were performed for every possible parameter combination (as defined above). These were carried out with and without temperature fluctuations. For the entire set of models for which fluctuations are not included, it was found that the ADF is larger than one 58% of the time. When larger than one, the ADF is between 1 and 2.03, with mean value 1.32±0.25. Nearly identical results are obtained when fluctuations are considered. Results isolating each parameter, with and without temperature and pressure fluctuations, are summarized in Table 1. The following conclusions can be established from this table:

(1) The abundance discrepancy factor does not depend on temperature and pressure fluctuations. Consequently, Peimbert's (1967) temperature "variations" should be interpreted as gradients in the ionized region.

(2) Whether ADF is greater or smaller than one, depends almost exclusively on the integration limits: it is always larger than 1 when $(t_1, t_2) = (0.75, 1.05)$, more than 75% of the time larger when $(t_1, t_2) = (0.85, 1.15)$ and always smaller when $(t_1, t_2) = (0.95, 1.25)$. Thus, the ADF cannot be larger than one when temperature in the ionized region is underestimated. This is not surprising. As noticed by Peimbert (1967), temperatures in ionized regions are overestimated since these are found from forbidden line ratios. These ratios are biased towards higher temperatures since, relative to the more intense line (such as [O III] 5007 Å), the weaker blue line (i.e., [O III] 4363 Å) will be predominantly produced in hotter regions.

(3) The pressure gradient has little effect on the ADF.

(4) Larger ADFs tend to be found in colder regions.

(5) Positive shallow temperature gradients ($\gamma = 2$) are less likely to produce ADFs larger than one. Negative temperature gradients favor larger ADFs.

(6) The largest values found for the ADF, always close to 2, are for $t_1 = 0.75$, $\gamma = -2$, $\beta = 0$ and $T_0 = 9000$. Except for the sign of the temperature gradient, it is possible that similar conditions are present in a number of H II regions (Esteban et al. 1998, 1999; García-Rojas et al. 2007).

TABLE 1

ADF FOR [O III] 5007 Å IN A SINGLE NEBULA, WITH AND WITHOUT FLUCTUATIONS

t_1, t_2	Without % ≥ 1	Mean ≥ 1	Mean ≤ 1	With % ≥ 1	Mean ≥ 1	Mean ≤ 1
0.75,1.05	100	1.48±0.21	-	100	1.48±0.20	-
0.85,1.15	75	1.11±0.07	0.97±0.02	79	1.11±0.08	0.97±0.02
0.95,1.25	0	-	0.83±0.06	0	-	0.84±0.07
β	Without % ≥ 1	Mean ≥ 1	Mean ≤ 1	With % ≥ 1	Mean ≥ 1	Mean ≤ 1
0.0	58	1.33±0.22	0.85±0.08	58	1.33±0.25	0.86±0.08
−0.5	58	1.32±0.23	0.86±0.08	61	1.30±0.23	0.86±0.08
T_0	Without % ≥ 1	Mean ≥ 1	Mean ≤ 1	With % ≥ 1	Mean ≥ 1	Mean ≤ 1
9000	58	1.44±0.32	0.81±0.09	58	1.44±0.31	0.81±0.09
12000	58	1.34±0.24	0.85±0.08	59	1.33±0.23	0.85±0.08
15000	58	1.28±0.19	0.88±0.06	60	1.27±0.18	0.88±0.06
18000	58	1.23±0.15	0.90±0.06	62	1.22±0.15	0.89±0.05
γ	Without % ≥ 1	Mean ≥ 1	Mean ≤ 1	With % ≥ 1	Mean ≥ 1	Mean ≤ 1
2	33	1.29±0.07	0.87±0.11	39	1.25±0.11	0.87±0.10
1	67	1.20±0.19	0.80±0.05	66	1.20±0.18	0.81±0.06
−1	67	1.35±0.25	0.86±0.03	67	1.35±0.24	0.87±0.03
−2	66	1.43±0.29	0.89±0.03	67	1.43±0.27	0.90±0.03

Numerical integrations were also carried out assuming that the CEL is [O III] 52 μm. For the entire set of models where fluctuations are not included, it was found that the ADF for the infrared line (ADF-IR) is larger than one 63% of the time. As expected, the abundance discrepancy factor was much smaller when FIR lines were considered: in this case the largest and mean values are 1.09 and 1.04 ± 0.03 respectively. At these temperatures, line emissivities from RLs and FIR CELs have a similar temperature response. As can be seen from Figures 1a,b,c,d, the integral terms for the RL and the CEL are nearly equal, so that the ADF-IR will be close to one as long as collisional de-excitation is inconsequential. But ADF-IR will be much larger than unity if the density is well above the critical value of [O III] 52 μm (~ 3500 cm^{-3} at 10^4 K), and this has not been acknowledged when line emissivity is determined.

Consequently, temperature gradients and fluctuations in a single ionized nebula cannot account for the mean ADF determined from optical CEL in H II regions and PNs, except under tight circumstances (see point 6 above). In objects where the density is smaller than a few thousand particles per cubic centimeter, the discrepancy between theory and observations is abysmal for FIR CELs. In the following section a more complex scenario is analyzed.

3. A COLLECTION OF IONIZED REGIONS

The extended nebula is now visualized as a collection of distinct ionized regions. At least one of these regions is as described in the preceding section. As a first order approximation the line intensity ratio is now given by,

$$\frac{I_r}{I_q} = \frac{I_r^0}{I_q^0}\,\frac{1+\sum_k I_r^k/I_r^0}{1+\sum_k I_q^k/I_q^0}\,. \tag{14}$$

Indices 0 and k stand for the ionized region described above and all the other additional regions included in the nebula (hereon, the primary and secondary components), and indices r and q stand for the recombination and collisionally excited lines. Following Equation (5) for either line intensity,

$$\frac{I^k}{I^0} = \frac{n_k(e)n_k(A^i)n_k(A)}{n_0(e)n_0(A^i)n_0(A)}\,\frac{P_k^2T_0^2r_k^3}{P_0^2T_k^2r_0^3}\,\frac{E(T_k)\int k}{E(T_0)\int 0}\,, \tag{15}$$

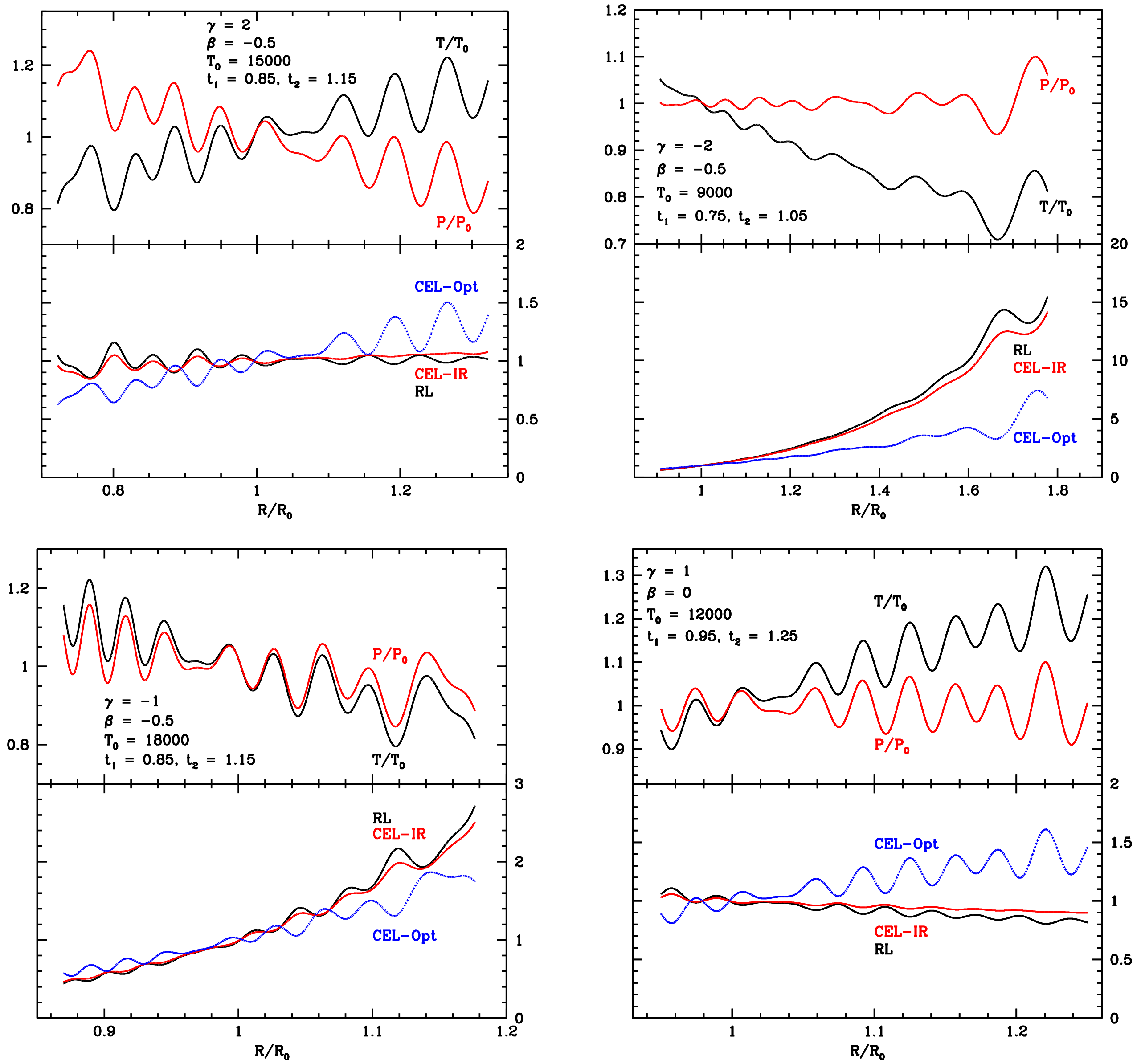

Fig. 1. Adimensional space (R/R_0) distributions for the adimensional temperature ($t = T/T_0$, in black) and pressure ($p = P/P_0$, in red) parameters, and the two integrands in Equation (11). RL stands for the recombination line emissivity integrand (in black), CEL-Opt for the [O III] 5007 Å emissivity integrand (in blue) and CEL-IR for the [O III] 52 μm emissivity integrand (in red).

with $\int k$ and $\int 0$ representing the integral term carried out over region k and the primary component, respectively.

The mass M_k contained in the region where $n_k(A^i) \approx 1$ is proportional to $r_k^3(\Delta r_k/r_k)m_k N_k$, where Δr_k is the width of the ring where $n_k(A^i) \approx 1$, and N_k and m_k are the mean particle density and mass per particle in this region. If all regions are composed of an ideal gas,

$$r_k^3 \; \alpha \; \frac{M_k}{(\Delta r_k/r_k) m_k P_k/T_k} \, . \qquad (16)$$

Doing the same for the primary component, the substitution of r_k^3 and r_0^3 in Equation (15) leads to

$$\frac{I^k}{I^0} = B_k(A^i)\ \frac{E(T_k)\int k}{E(T_0)\int 0}\,, \tag{17}$$

with $B_k(A^i)$ defined as

$$B_k(A^i) = \frac{n_k(e)n_k(A^i)n_k(A)}{n_0(e)n_0(A^i)n_0(A)}\ \frac{\Delta r_0/r_0}{\Delta r_k/r_k}\ \frac{T_0 P_k M_k m_0}{T_k P_0 M_0 m_k}\,, \tag{18}$$

where M_0 and m_0 are the mass and mean mass per particle of the region in the primary component where $n_0(A^i) \approx 1$. Consequently,

$$\mathrm{ADF} = \frac{\int R(0)}{\int Q(0)}\ \frac{1+\sum_k B_k(A^i)\ R(T_k)/R(T_0)\ \int R(k)/\int R(0)}{1+\sum_k B_k(A^i)\ Q(T_k)/Q(T_0)\ \int Q(k)/\int Q(0)}\,, \tag{19}$$

where $\int R(0)$ and $\int Q(0)$ are the integrals in the numerator and denominator of the RHS of Equation (6), and $\int R(k)$ and $\int Q(k)$ are these integrals applied to region k.

To facilitate analysis, the following simplifying assumptions for the secondary component will be helpful:

(1) All regions are uniform, that is, electron density, ion concentration, abundance, pressure and temperature are constant in the region where $n_k(A^i) \approx 1$. Thus, $\int R(k) = \int Q(k) = 1$.

(2) All regions are identical.

Following Equation (19),

$$\mathrm{ADF} = \frac{\int R(0)}{\int Q(0)}\ \frac{1+\aleph B(A^i)\ R(T_c)/[R(T_0)/\int R(0)]}{1+\aleph B(A^i)\ Q(T_c)/[Q(T_0)/\int Q(0)]}\,, \tag{20}$$

where $B(A^i) = B_k(A^i)$ and $T_c = T_k\ \ \forall\ k$, and $\aleph$ is the number of regions in the secondary component. Notice that Equation (20) is identical to Equation (6) when $\aleph B(A^i) = 0$, i.e., when the secondary component is absent.

Under this scenario the underlying cause of the abundance discrepancy is a temperature difference between the primary and secondary components. It is clear that the ADF can only be larger than one when

$$\frac{R(T_c)Q(T_0)}{R(T_0)Q(T_c)} > \frac{\int R(0)}{\int Q(0)}\,. \tag{21}$$

The RHS of this inequality has already been computed (Equation 11), and the left hand side (hereon LHS) is plotted in Figure 2 as a function of $t_c = T_c/T_0$, for $T_0 = 9000$, 12000, 15000 and 18000 K and [O III] 5007 Å. As can be seen, $R(T_c)Q(T_0)/R(T_0)Q(T_c)$ increases sharply when $t_c < 1$. Thus, if the inequality is valid at a certain t_c, it will be more pronounced when this quantity is smaller. Consequently, the secondary component must be much colder than the primary component when the ADF is large. Obviously, this implies that the abundance discrepancy effect cannot be produced by external heating sources. Notice too that if the inequality holds, the ADF will increase with $\aleph B(A^i)$.

Equation (20) is used to determine the value of $\aleph B(A^i)$ for which ADF = 2, 5, 10 and 20, when the CEL is [O III] 5007. Calculations were performed for the physical scenarios considered in § 2, and situations where temperature in the secondary component was less than the smallest temperature $T_1/T_0 = t_1$ in the primary component (notwithstanding fluctuations), i.e., for $T_c/T_1 = t_c/t_1 = 0.95$, 0.75, 0.5, 0.25, 0.1 and 0.05. A success ratio is defined as the fraction of cases where the prescribed ADF is found under the condition that $\aleph B(A^i) \leq 10000$.

Grouping all models for the primary component, Table 2 lists the success ratio as a function of the temperature contrast T_c/T_0. Conditions to produce ADF = 2 were found possible even for a relatively mild temperature contrast ($T_c/T_0 = 0.81$), but larger ADFs are only viable when there is a very significant temperature difference. For instance, an ADF > 5 implies that temperature in the hot primary component must be more than twice as large as in the cold component. All other parameters, such as the shape of the pressure and temperature gradients in the primary component, have a minor effect on the outcome and practically none on the success ratio.

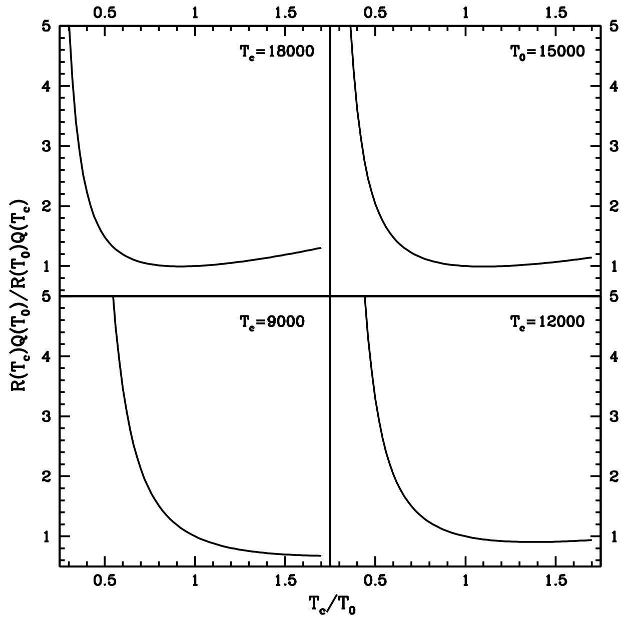

Fig. 2. $R(T_c)Q(T_0)/R(T_0)Q(T_c)$ as a function of the temperature contrast $t_c = T_c/T_0$, for $T_0 = 9000$, 12000, 15000 and 18000 K and [O III] 5007Å. Inequality (21) holds if this product is large.

TABLE 2

SUCCESS RATIO FOR A TWO COMPONENT NEBULA AS A FUNCION OF THE TEMPERATURE CONTRAST T_C/T_0

T_c/T_0	ADF=2	ADF=5	ADF=10	ADF=20
0.90	0.00	0.00	0.00	0.00
0.81	0.25	0.00	0.00	0.00
0.71	1.00	0.00	0.00	0.00
0.64	1.00	0.25	0.00	0.00
0.56	1.00	0.75	0.25	0.00
0.47	1.00	1.00	0.75	0.25
0.42	1.00	1.00	1.00	0.50
0.38	1.00	1.00	1.00	1.00
0.04	1.00	1.00	1.00	1.00

Equation (20) is now used to determine the ADF when the CEL is [O III] 52μm, using the value of $\aleph B(A^i)$ for which the ADF = 2, 5, 10 and 20 when the CEL is [O III] 5007 Å. The ADF for the far-infrared line, ADF-IR, is plotted as a function of the temperature contrast $t_c = T_c/T_0$ in Figure 3. Cases where the optical ADF (ADF-Opt) is equal to 2, 5, 10 and 20 are shown as black crosses, red empty circles, green empty squares and blue plus signs, respectively. It is plain to see that ADF-Opt is larger than ADF-IR under nearly all circumstances, and in the mean ADF-IR = 1.47±0.60, 1.93±0.60, 2.25±0.78 and 2.58±0.94 when ADF-Opt is equal to 2, 5, 10 and 20 respectively. This makes physical sense and is supported by observations. As can be seen in Figure 4, ADF-IR is usually smaller than ADF-Opt (data for O^{+2} from Tsamis et al. 2004).

But Figure 3 also shows that temperature in the cold region must be much smaller than in the primary component if ADF-IR is large. For instance, ADF-IR $\geq$ ADF-Opt ≈ 2 only when $T_c/T_0 \leq 0.05$, i.e., when the temperature in the secondary component of the ionized region is ~ 500 K. Smaller temperatures are required to

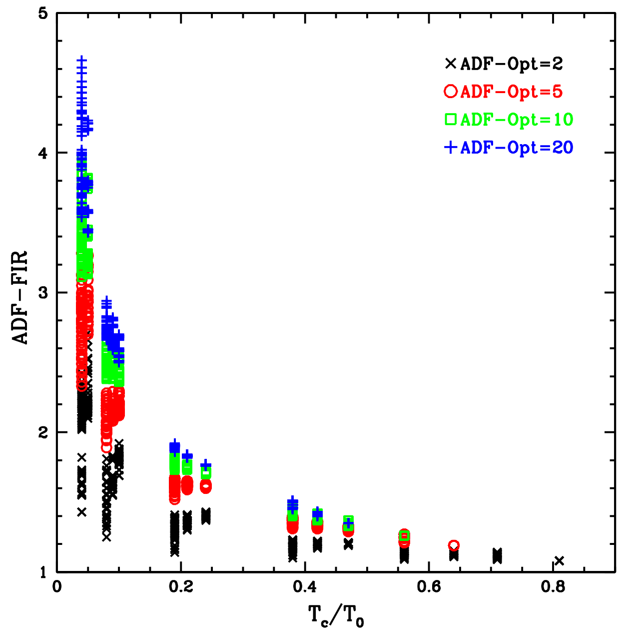

Fig. 3. The ADF for the [O III] 52 μm line (ADF-IR) as a function of the temperature ratio $t_c = T_c/T_0$ using the same set of parameters required to produce an ADF = 2 (black crosss), 5 (red empty circles), 10 (green empty squares) and 20 (blue plus signs) for [O III] 5007 Å. Model marks for ADF = 20, 10 and 5 may be superimposed on model marks for ADF = 10, 5 and 2, 5 and 2, and 2, respectively.

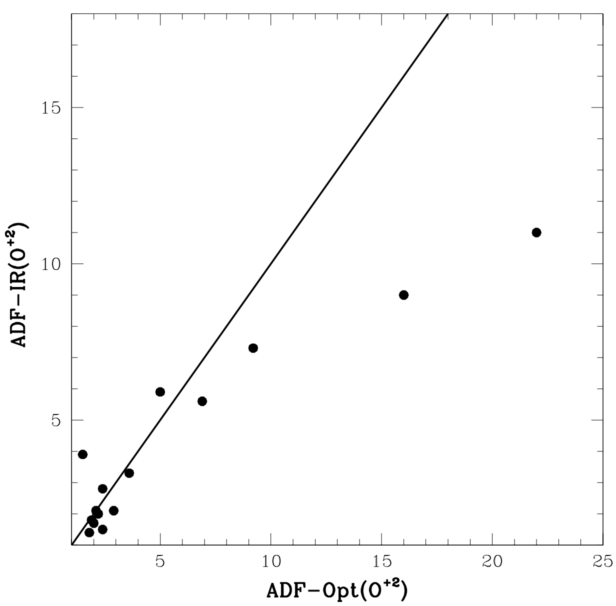

Fig. 4. The ADF for infrared O^{+2} CELs (ADF-IR) *vs.* the ADF for optical CELs from the same ion (ADF-Opt). The entire data set is from Tsamis et al. (2004). The straight line is for ADF-Opt = ADF-IR.

match larger ADFs: T_c must be close to 200, 100 and 50 K if ADF-Opt $\sim$ ADF-IR $\sim$ 5, 10 and 20 respectively. It is hard to imagine how a stable ionized region can be so cold, though temperatures close to, or smaller than, 500 K have been found in three PNs (with ADFs > 5) from O II 4089/4649 (Tsamis et al. 2004; Liu et al. 2006). On the other hand, large values for ADF-IR may be associated to ignoring collisional de-excitation of the fine structure IR lines, since at 500 and 100 K the density in the cold region is probably larger than the

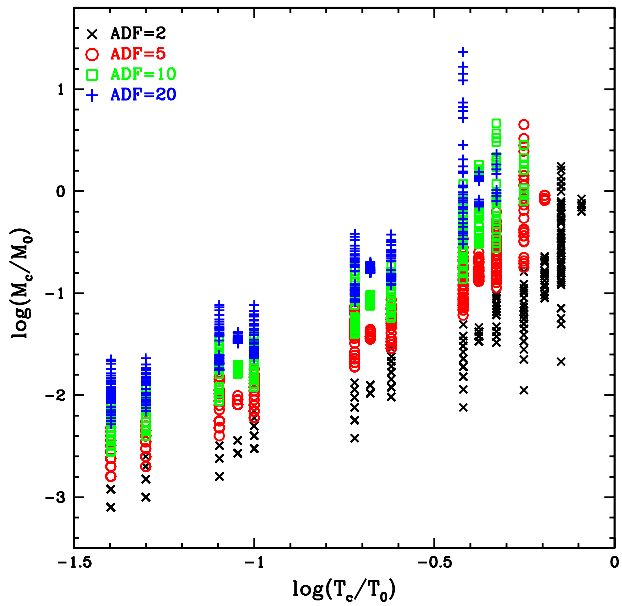

Fig. 5. Logarithm of the mass ratio $[M_c/M_0]_{A^i}$ as a function of the logarithm of the temperature ratio $t_c = T_c/T_0$ for [O III] 5007 Å and ADF = 2 (black crosss), 5 (red empty circles), 10 (green empty squares) and 20 (blue plus signs). Model marks for ADF = 20, 10 and 5 may be superimposed on model marks for ADF = 10, 5 and 2, 5 and 2, and 2, respectively.

critical value for [O III] 52μm (1400 and 5600 cm^{-3}), but possibly not if T = 50 K, since the critical density is 63000 cm^{-3} at this unimaginable temperature for an ionized region.

4. CONSEQUENCES OF A COLD IONIZED COMPONENT

The existence of a cold ionized component may have important consequences on the way data and observations of H II regions and PNs are interpreted. From Equation (18) it can be deduced that the out-coming value of $\aleph B(A^i)$ is proportional to the mass of the secondary component, M_c, relative to the mass M_0 of the hot region. The secondary-to-primary component mass ratio in regions where $n_c(A^i)$ and $n_0(A^i) \simeq 1$ is given by

$$\left[\frac{M_c}{M_0}\right]_{A^i} = \frac{\Delta r_c/r_c}{\Delta r_0/r_0}\frac{P_0}{P_c}\frac{n_0(A)m_c}{n_c(A)m_0}\aleph B(A^i)\ t_c\,, \tag{22}$$

since $n_0(e)$ and $n_c(e)$, the electron-to-total particle density ratios, are both close to 0.5. The relative width of regions where $n_c(A^i)$ and $n_0(A^i) \simeq 1$ depends on the heating and cooling rates within them. Both are illuminated by the same star, but differences in other influential parameters, such as chemical composition or gas-to-dust ratio, are still to be found. Thus, it is conservatively assumed that $(\Delta r_c/r_c)/(\Delta r_0/r_0) \approx 1$. The secondary component will expand, cool and de-pressurize (in an adiabatic process) as long as $P_c > P_0$. To preserve this hypothetical structure, pressure in the hot component must be slightly larger than in the cold region. The two remaining factors depend on the chemical composition of each component. In the case of PNs it has been suggested that the abundance of heavy elements relative to helium plus hydrogen is constant, but that He/H is considerably larger in the cold component (Liu et al. 2000). If this is so, $n_0(A)/n_c(A) \simeq 1$, but m_c/m_0 and $n_0(H)/n_c(H) \approx 2$ in the extreme case where hydrogen has been depleted by a factor of 10 (N_c(H) = N_c(He)). In H II regions the situation is somewhat more uncertain, if these have been contaminated by exogenous material (Tsamis & Pequignot 2005; Stasińska et al. 2007).

Consequently, if this hypothetical model for ionized nebulae is valid and these assumptions are adequate, the cold-to-hot mass ratio in regions where A^i is predominant is approximately given by (maybe twice as large in some PNs),

$$\left[\frac{M_c}{M_0}\right]_{A^i} \simeq \aleph B(A^i)\ t_c\,. \tag{23}$$

This relationship will hold for any ion if the preceding conditions are fulfilled. If this is so, $\aleph B(X^j)/\aleph B(A^i) \simeq [M_c/M_0]_{X^j}/[M_c/M_o]_{A^i}$ for any pair of ions A^i and X^j. A quantity proportional to the mass ratio was used by Stasińska (2002) to define the relative weights of the emitting regions in a two temperature component "toy" model. This quantity is a free parameter in Stasińska's model, whereas here it depends on the ADF ant the temperature ratio T_c/T_0, both observable quantities.

The mass ratio $(M_c/M_0)_{A^i}$ as a function of the temperature contrast T_c/T_0 is plotted in Figure 5 (log-log) for all model outcomes where the CEL is [O III] 5007 Å and ADF = 2 (black crosses), 5 (red empty circles), 10 (green empty squares) and 20 (blue plus signs). The implications that can be derived from this figure are self-evident and sensible. For a fixed temperature contrast T_c/T_0, the ADF can only be increased by adding cold matter to the region. Secondly, a smaller amount of cold matter is required to produce a given ADF, as temperature in the cold region is increasingly smaller than in the primary component. Furthermore, though there is a wide spread of values for M_c/M_0 for a fixed temperature contrast, it is clear that $M_c/M_o \; \alpha \; (T_c/T_0)^\beta$ for any ADF. Mean square fits applied to the mean values of the mass ratio at each T_c/T_0 imply that β is between 1.8 (for ADF = 2) and 2.1 (for ADF = 20), with correlation coefficients exceeding 0.97 in all cases (but large typical errors reflecting the large spread of values for M_c/M_0). Finally, and significantly, in nearly all cases the model predicts that there will be more matter in the hot primary component than in the cold region, hundreds of times more when $T_c/T_0 \approx 0.05$, the temperature contrast where ADF-Opt $\simeq$ ADF-IR $\approx$ 2. Thus, an accurate determination of abundances in the hot primary component is particularly important to model stellar structure and the chemical evolution of galaxies and the universe, among other things.

Notice that a small amount of cold matter inclusions can have a very significant effect on the spectral properties of the ionized region. This result is diametrically opposed to Mathis'(1995) findings from a basic two temperature component model, where a small amount of hot gas has more effect than a large amount of cold material, from which he concludes that it would be difficult to explain the ADF with cool spots. Using a different formalism for the basic two temperature component model, Stasińska (2002) is also inclined against cool spots, and favors a model where the temperature contrast is produced by external heating sources, such as shock and heat conduction fronts, acting on a small number of regions. The discovery of very low temperatures from RL ratios (Tsamis et al. 2004; Liu et al. 2006) supports the results presented in this paper, but not the conclusions of Mathis (1995) and Stasińska (2002).

Mass estimates must be reconsidered if there is a cold component in addition to the hot region where optical CELs shine brightly. Assuming that $[M_c/M_0]_{\mathrm{H}^+} = [M_c/M_o]_{A^i}$, the mass of ionized hydrogen, $M(\mathrm{H}^+) = M_0(\mathrm{H}^+) + M_c(\mathrm{H}^+)$, can be found from

$$M(\mathrm{H}^+) \simeq M_0(\mathrm{H}^+) \; \left[1 + \aleph B(A^i) \; t_c\right] \; . \tag{24}$$

If the mass of ionized hydrogen is found from the intensity of a hydrogen recombination line, $I(\mathrm{H}^+) = I_0(\mathrm{H}^+) + I_c(\mathrm{H}^+)$, following the preceding assumptions it can be shown that

$$M(\mathrm{H}^+) = \frac{m_H I(\mathrm{H}^+)}{N_0(e) R(\mathrm{H}^+, T_0)} \; \mathrm{MCF} \, , \tag{25}$$

where $N_0(e)$ is the electron density in the hot primary component (usually determined from [S II]6717/[S II]6731), $R(\mathrm{H}^+, T_0)$ is the line emissivity of H^+ (following Osterbrock & Ferland 2006, it is proportional to $T_0^{-0.83}$), and

$$\mathrm{MCF} = \frac{\int hh^+(t^{\gamma\beta} + \delta p) t^{3\gamma - 1}/(t + \delta t) \; dt}{\int \mathbf{\Gamma} \; (t + \delta t)^{-0.83} \; dt} \; \frac{1 + \aleph B(A^i) \; t_c}{1 + \mathrm{BIC}} \, , \tag{26}$$

with $\mathbf{\Gamma}$ defined in Equation (12). BIC stands for the Balmer line intensity contrast between the cold and hot components of the ionized region, $I_c(\mathrm{H}^+)/I_0(\mathrm{H}^+)$. Following Equations (5) and (17) and all the preceding assumptions

$$\mathrm{BIC} = \frac{\aleph B(\mathrm{H}^+) \; t_c^{-0.83}}{\int \mathbf{\Gamma} \; (t + \delta t)^{-0.83} \; dt} \, . \tag{27}$$

Since it has been assumed that $[M_c/M_0]_{\mathrm{H}^+} = [M_c/M_o]_{A^i}$, $\aleph B(\mathrm{H}^+) = \aleph B(A^i)$.

The first term on the RHS of Equation (25), $m_H I(\mathrm{H}^+)/N_0(e)R(\mathrm{H}^+, T_0)$, is the mass of ionized hydrogen in a region where density and temperature are constant. This is the usual way mass estimates are determined from the intensity of a hydrogen recombination line. Thus, MCF is a mass correction factor that takes into account more complex situations, including the possible presence of a cold component that is practically invisible in optical CEL emission, but bright when seen in the light of recombination lines, including those produced by hydrogen.

Since the emissivity of a hydrogen recombination line is larger at smaller temperatures, it follows that fewer hydrogen atoms are required to produce a given line intensity when matter is cold. In consequence, hydrogen masses will be overestimated if there is cold ionized gas and it's presence is not acknowledged. This reasoning is captured in Figure 6, where all model outcomes for the MCF as a function of the temperature contrast $t_c = T_c/T_0$, are plotted. As before, models for which ADF = 2, 5, 10 and 20 are shown as black crosses, red empty circles, green empty squares and blue plus signs, respectively (models are for [O III] 5007 Å). Though there is a wide spread of results, it is clear that the MCF is smaller for large ADFs. In the mean, MCF = 0.4, 0.2, 0.1 and 0.06 when ADF = 2, 5, 10 and 20. Thus, H^+ mass estimates of H II regions and PNs based on Balmer line intensities may have been overestimated between 2 and 10 times, if the observed ADFs are indeed disclosing the existence of a cold ionized component in these emission nebulae.

The logarithm of the Balmer line intensity contrast between the secondary and primary components of the ionized region, BIC, is plotted in Figure 7 for [O III] 5007 Å and ADF = 2 (black crosses), 5 (red empty circles), 10 (green empty squares) and 20 (blue plus signs). There is a large spread of values for the BIC, down from 0.1 and up to 257, but the cold component is brighter under most circumstances. The line intensity contrast clearly tends to increase with larger ADFs and, in the mean, the line intensity in the cold component is 2 to 24 times brighter as the ADF changes from 2 to 20. This makes physical sense, since Balmer line emissivity is inversely proportional to the electron temperature, but it does not appear to be consistent with observations carried out by Tsamis et al. (2008) in NGC 6153, where they find that regions of low Hβ surface brightness exhibit the largest ADFs. Notice too that, if this scenario is correct, line ratio images such as [O III]5007/Hα, cannot be interpreted as exclusively depicting variations in the excitation conditions of a single component nebula, but also as maps of the relative distribution of hot and cold gas. Even more consequential is that the brightest Balmer line component of the ionized region usually occurs where there is less mass, an additional complication to the challenging problem of determining accurate abundances in the usually hotter, fainter (in H^+ and other recombination lines) and more massive component of the nebula.

Astronomical instrumentation cannot separate light from the hot and cold components, but abundances are still being calculated as if this were so even when this scenario is being explicitly assumed. At its most basic level, the two component model can be visualized as two uniform regions, "0" and "c", coexisting in an emission nebula excited by a single star. As mentioned before, simple versions of the two temperature component model, each using a different formalism, have been put forward by Mathis (1995) and Stasińska (2002). The starting point in this paper is that any line intensity is given by,

$$I(A^i) = N_0(A^i)N_0(e)E(A^i, T_0)V_0(A^i) + N_c(A^i)N_c(e)E(A^i, T_c)V_c(A^i)\,, \tag{28}$$

where $V_0(A^i)$ and $V_c(A^i)$ are the volumes occupied by regions "0" and "c" where A^i is the prevalent ion. Taking into account that $V = M/Nm$ —where M, N and m are the total mass, the total density and the mass per particle— and assuming that $N_0(e)/N_0 = N_c(e)/N_c$, pressure is uniform ($P_c = P_0$) and $M_c(A^i)/M_0(A^i) = (T_c/T_0)^\beta$ (see Figure 5), the line intensity can now be re-written as

$$I(A^i) = N_0(A^i)N_0(e)E(A^i, T_0)V_0(A^i)\left[1 + \Lambda_{A^i}\left[\frac{T_c}{T_0}\right]^{\beta-1}\frac{E(A^i, T_c)}{E(A^i, T_0)}\right]\,, \tag{29}$$

where $\Lambda_{A^i} = [m_0/m_c][N_c(A^i)/N_c]/[N_0(A_i)/N_0]$ is the parameter defining chemical differences between the two components. In this simple model the abundance discrepancy factor is given by

$$\mathrm{ADF} = \frac{1 + \Lambda_{A^i}(T_c/T_0)^{\beta-1}R(A^i, T_c)/R(A^i, T_0)}{1 + \Lambda_{A^i}(T_c/T_0)^{\beta-1}Q(A^i, Tc)/Q(A^i, T_0)}\,. \tag{30}$$

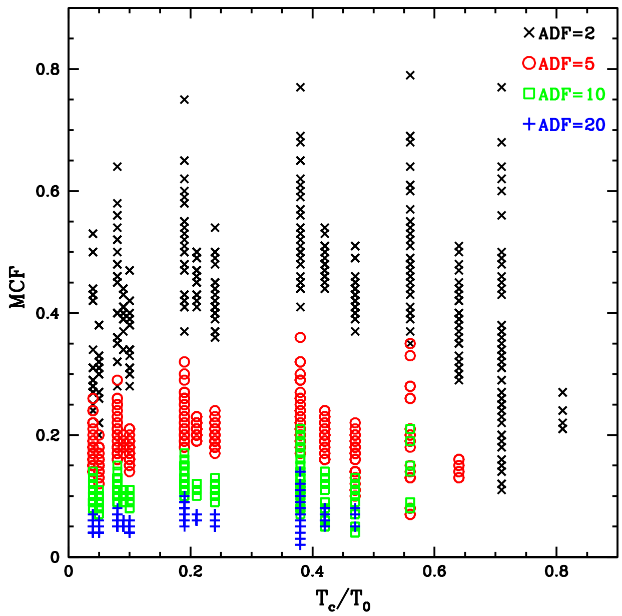

Fig. 6. Mass correction factor (MCF) as a function of the temperature ratio $t_c = T_c/T_0$ for [O III] 5007 Å and ADF = 2 (black crosss), 5 (red empty circles), 10 (green empty squares) and 20 (blue plus signs). Model marks for ADF = 20, 10 and 5 may be superimposed on model marks for ADF = 10, 5 and 2, 5 and 2, and 2, respectively.

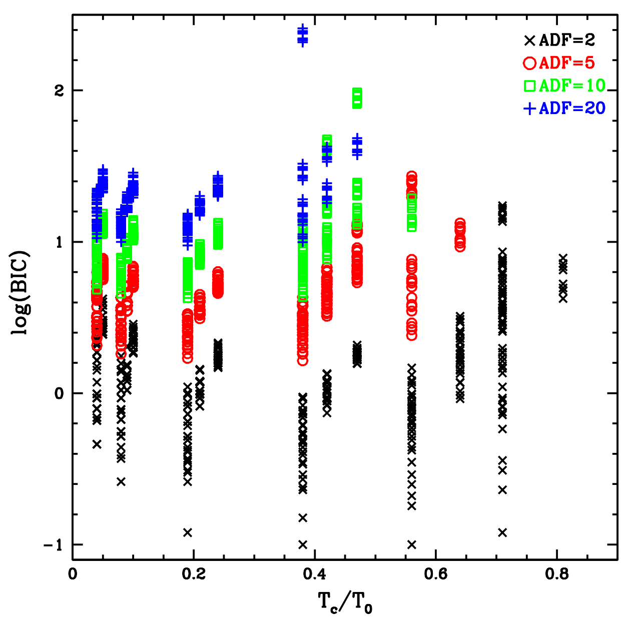

Fig. 7. Logarithm of the Balmer line intensity contrast (BIC) as a function of the temperature ratio $t_c = T_c/T_0$ for [O III] 5007 Å and ADF = 2 (black crosss), 5 (red empty circles), 10 (green empty squares) and 20 (blue plus signs). Model marks for ADF = 20, 10 and 5 may be superimposed on model marks for ADF = 10, 5 and 2, 5 and 2, and 2, respectively.

Equation (30) is plotted in Figure 8 for $\beta = 2, 1.5, 1$ and 0.5. The graph is for $\Lambda_{A^i} = 0.5$, which could correspond to severe hydrogen depletion in the cold component, a scenario that has been regarded as possible in some PN. Tsamis et al. (2003, 2004) and Liu et al. (2006) determined the ADF and various electron temperatures for a few PNs. Temperature ratios derived from these works are included in this figure: crosses

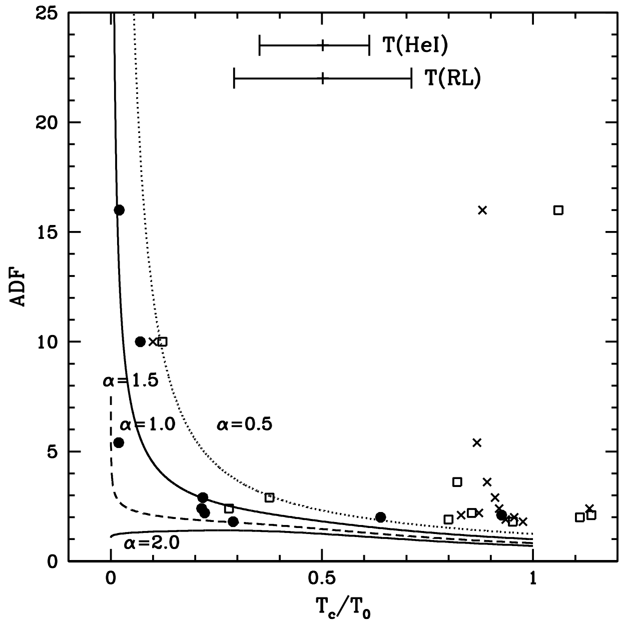

Fig. 8. Abundance discrepancy factor (ADF) as a function of the temperature ratio $t_c = T_c/T_0$ as given by Equation (30), for $\beta = 2, 1.5, 1$ and 0.5. Data points are from Tsamis et al. (2003, 2004) and Liu et al. (2006). T_e(O II 4089/4649)/T_e([O III] 5007/4363) is represented with full circles, crosses stand for T_e(BJ/H11)/T_e([O III] 5007/4363), where BJ is the Balmer jump, and empty squares represent T_e(He I 6678/4471)/T_e([O III] 5007/4363). Mean temperature ratio error bars for T_e(O II 4089/4649)/T_e([O III] 5007/4363) and T_e(He I 6678/4471)/T_e([O III] 5007/4363) are named T(RL) and T(He I) respectively, and were deduced from the work of Tsamis et al. (2004). The ADF for PN Hf 2-2 was determined from the work of Liu et al. (2006), recalculating the RL abundance of O^{+2} using T_e([O III] 5007/4363.

for T_e(BJ/H11)/T_e([O III] 5007/4363), full circles for T_e(O II 4089/4649)/T_e([O III] 5007/4363) and empty squares for T_e(He I 6678/4471)/T_e([O III] 5007/4363). As can be seen, there is no match between any one of the curves derived from Eq.(30) and the ratio including the Balmer jump temperature, T_e(BJ/H11)/T_e([O III] 5007/4363). But there is an indication that the model is consistent with observations if $\beta \simeq 1$ and either one of the other two temperature ratios, particularly the one involving the O II recombination lines, is associated with T_c/T_0. This is quite satisfactory, since temperatures in the cold and hot components are naturally identified with T_e(O II 4089/4649) and T_e([O III] 5007/4363). Unfortunately, there is no match when $\beta = 2$, the exponent that best fits the $M_c/M_0 = (T_c/T_0)^\beta$ relationship found in a more complex scenario (see Figure 5). It is worth noting that the shape of the curve does not change when different values are used for Λ_{A^i}, underscoring the fact that abundance discrepancies are effected by temperature contrasts, regardless of their origin.

Notice too that if $V_0(\mathrm{H}^+) \sim V_0(A^i)$, the abundance of any ion A_i with respect to H^+ in one of the components of this model nebula is

$$\frac{N_0(A^i)}{N_0(\mathrm{H}^+)} = \frac{I(A^i)R(\mathrm{H}^+, T_0)}{I(\mathrm{H}^+)E(A^i, T_0)} \frac{1 + \Lambda_{\mathrm{H}^+}(T_c/T_0)^{\beta-1} R(\mathrm{H}^+, T_c)/R(\mathrm{H}^+, T_0)}{1 + \Lambda_{A^i}(T_c/T_0)^{\beta-1} E(A^i, T_c)/E(A^i, T_0)} \,. \tag{31}$$

But it cannot be found directly without knowing in advance the chemical differences between the two components, Λ_{A^i} and Λ_{H^+}. Notice that Equation (30) can be used to find Λ_{A^i} from the observed values of the ADF, T_0 (T_e([O III] 5007/4363)) and T_c (T_e(O II 4089/4649)). But Λ_{H^+} cannot be determined without establishing the abundance ratio for the two components, since this parameter is equal to $\Lambda_{A^i}[N_0(A^i)/N_0(\mathrm{H}^+)]/[N_c(A^i)/N_c(\mathrm{H}^+)]$. It is clear that additional information is required to establish a reliable method to calculate abundances in a two temperature component nebula. The solution may well have to wait for a model that can explain how substantial amounts of cold, dense ionized gas can survive in the high energy density environment created by the exciting stars of H II regions and PNs.

5. CONCLUSIONS

A single ionized region with temperature gradients will produce an ADF based on an optical CEL (ADF-Opt) larger than one only when the mean temperature is not underestimated. Peimbert's (1967) temperature variations must be interpreted as gradients, since fluctuations have almost no effect. When larger than one, the mean value of ADF-Opt is 1.3, significantly less than the mean found in H II regions and PNs. The largest ADF-Opt that can be found is close to 2, and is always associated to situations where the mean temperature is on the low side ($T_0 \sim 9000$ K) and overestimated (as expected). Such a region will not yield an ADF based on a FIR CEL (ADF-IR) larger than 1.1, much less than observed. Higher values of ADF-IR may occur when the density is larger than the critical value for these IR lines, and this is not taken into account when line emissivity is calculated. Thus, single ionized regions usually fail to account for the observed abundance deficiency factors.

If additional components are added to the emission region, these must be colder in order to explain the observed ADFs. This implies that the abundance discrepancy cannot be associated to heating mechanisms, such as shock and conduction fronts. Conditions to produce ADF = 2 are possible even when there is a mild temperature contrast between the "hot" and "cold" components ($T_c/T_0 = 0.81$), but larger ADFs can only be reproduced when there is a substantial temperature difference (for instance, ADF-Opt ≥ 5 only when $T_c/T_0 < 0.5$). As expected and observed, ADF-Opt is larger than ADF-IR under nearly all circumstances. Large values of ADF-IR require very low temperatures in the cold component, i.e., about 500, 200, 100 and 50 K if ADF-IR $\simeq$ 2, 5, 10 and 20 and $T_0 \sim 10^4$ K. Temperatures close to 500 K have been measured in a few PNs (Tsamis et al. 2004; Liu et al. 2006). But great care should be taken when the ADF-IR is analyzed, since collisional de-excitation of these fine structure IR lines can be quite significant in the cold and dense component of the emission nebula. If this is not taken into account, the emissivity and the ADF will be overestimated.

The relative amount of mass in the hot (M_0) and cold (M_c) components also has an important effect. The cold secondary component may be relatively warm and heavy or cold and light, and still produce the same ADF. It is shown that, for any any given ADF, M_c/M_0 α $(T_c/T_0)^\beta$, with $\beta \approx 2$. In nearly all cases there is more matter in the hot region, hundreds of times more when the secondary component is particularly cold. This implies that a small amount of cold matter will have a large effect on the emission spectrum. It also implies that an accurate determination of abundances in the hot primary component is particularly important, since most of the mass is in this region.

Mass estimates of ionized hydrogen based on the intensity of a Balmer line must be reconsidered if there is an additional cold component. It is found that these may have been overestimated between 2 and 10 times, with the largest discrepancies being associated to larger ADFs. It is also shown that under nearly all circumstances, Balmer lines should be 2 to over 20 times brighter in the cold component, as ADF-Opt changes from 2 to 20. Consequently, line ratio images such as [O IIII]5007/Hβ, may render the relative distribution of cold and hot matter in addition to variations in the excitation conditions.

Observations are consistent with a simple two temperature component model relating the abundance discrepancy factor and the temperature contrast T_c/T_0, if it is assumed that M_c/M_0 α (T_c/T_0), and T_e(O II 4089/4649) and T_e([O III] 5007/4363) are identified with the temperatures of the cold and hot components. It is shown that there is insufficient information to determine the chemical composition of the two components using this simple model.

The determination of accurate abundances with a two component model is hindered by the fact that observations cannot discriminate light from the hot and cold gas. Additional circumstances weigh against this goal. Since there is only one heating source, cooling must be more efficient in the colder region. This implies that heavy elements must be more abundant in the gaseous phase of this component, particularly when density is so large that collisional de-excitation inhibits radiative cooling by fine structure IR lines. Thus, at least two separate and probably quite different abundance sets must be quantified. On the other hand, most of the mass is usually contained in the hot component dominated by optical CEL emission, but Balmer lines are generally more intense in the cold region. Thus, spectral information is incomplete for the region where it is more needed. All these suggests that, at the present time, it does not seem possible to have little more than an educated guess or an order of magnitude estimate of the abundance of heavy elements when there is an additional cold component and the ADF is larger than ~ 2.

The two component model, an extreme version of the disputed temperature fluctuations proposed by Peimbert (1967), may provide a solution to the problem of abundance discrepancies derived from CELs and RLs, but it has some disturbing consequences on the way emission nebulae are interpreted, and opens important and difficult questions. To start with, since light from the cold and hot regions is blended, it is not clear how reliable chemical compositions for the two components can be found. More importantly, the credibility of this scenario depends on the construction of a plausible model for the origin and permanence of cold (or very cold!) and dense (maybe too dense !) ionized regions, in a medium that has to be several times hotter. On the other hand, if the basic physics are essentially right, it does not seem possible to explain the existence of the ADF without arguing for large temperature contrasts within emission nebulae.

Comments from an anonymous referee and support from DGAPA-UNAM project IN102607-3 are gratefully acknowledged.

REFERENCES

Esteban, C., Peimbert, M., Torres-Peimbert, & Escalante, V. 1998, MNRAS, 295, 401

Esteban, C., Peimbert, M., Torres-Peimbert, S., García-Rojas, J., & Rodríguez, M. 1999, ApJS, 120, 113

García-Rojas, J., Esteban, C., Peimbert, A., Rodríguez, M., Peimbert, M., & Ruiz, M. T. 2007, RevMexAA, 43, 3

Garnett, D. R., & Dinerstein, H. L. 2001a, RevMexAA (SC), 10, 13

__________. 2001b, ApJ, 558, 145

Grandi, S. A. 1976, ApJ, 206, 658

Kholtygin, A. F., & Feklistova, T. 1992, AZh, 69, 960

Liu, X.-W., Barlow, M. J., Zhang, Y., Bastin, R. J., & Storey, P. J. 2006, MNRAS, 368, 1959

Liu, X.-W., Storey, P. J., Barlow, M. J., & Clegg, R. E. S. 1995, MNRAS, 272, 369

Liu, X.-W., Storey, P. J., Barlow, M. J., Danziger, I. J., Cohen, M., & Bryce, M. 2000, MNRAS, 312, 585

Liu, Y., Liu, X.-W., Barlow, M. J., & Luo, S.-G. 2004, MNRAS, 353, 1251

Mathis, J. S. 1995, RevMexAA (SC), 3, 207

Marigo, P. 2001, A&A, 370, 194

Nikitin, A. A., Kholtygin, A. F., Sapar, A., & Feklistova, T. 1994, Baltic Astronomy, 3, 112

Osterbrock, D. E., & Ferland, G. J. 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei (2nd ed.; Sausalito: University Science Books)

Peimbert, M. 1967, ApJ, 150, 825

Peimbert, M., Storey, P. J., & Torres-Peimbert, S. 1993, ApJ, 414, 626

Robertson-Tessi, M., & Garnett, D. R. 2005, ApJS, 157, 371

Stasińska, G. 2002, RevMexAA (SC), 12, 62

Stasińska, G., Tenorio-Tagle, G., Rodríguez, M., & Henney, W. J. 2007, A&A, 471, 193

Storey, P. J. 1994, A&A, 282, 999

Tsamis, Y. G., Barlow, M. J., Liu, X.-W., Danziger, I. J., & Storey, P. J. 2003, MNRAS, 345, 186

Tsamis, Y. G., Barlow, M. J., Liu, X.-W., Storey, P. J., & Danziger, I. J. 2004, MNRAS, 353, 953

Tsamis, Y. G., & Péquignot, D. 2005, MNRAS, 364, 687

Tsamis, Y. G., Walsh, J. R., Péquignot, D., Barlow, M. J., & Danziger, I. J. 2008, MNRAS, 386, 22

van den Hoek, L. B., & Groenewegen, M. A. T. 1997, A&AS, 123, 305

Zatsarinny, O., Gorczyca, T. W., Korista, K. T., Badnell, N. R., & Savin, D. W. 2004, A&A, 417, 1173

J. Bohigas: Instituto de Astronomía, Universidad Nacional Autónoma de México, Apdo. Postal 877, 22830 Ensenada, B. C., Mexico (jbb@astrosen.unam.mx).