A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
3(03.13.2;03.13.6;08.06.2;10.19.3;11.09.4;11.19.2)

ASTRONOMY AND ASTROPHYSICS
19.5.1997

The star formation law in disk galaxies — a Bayesian view

J. Köppen[1,3,4] **and H.-E. Fröhlich**[2]

[1] Institut für Astronomie und Astrophysik der Universität, D–24098 Kiel, Germany
[2] Astrophysikalisches Institut Potsdam, An der Sternwarte 16, D–14482 Potsdam, Germany
[3] Observatoire Astronomique, 11 Rue de l'Université, F–67 000 Strasbourg, France
[4] International Space University, Parc d'Innovation, F–67400 Illkirch, France

May 19, 1997

Abstract. The dependence of the star formation rate $\Psi \propto g^x/r^y$ on gas surface density g and galactocentric radius r are studied in disk galaxies. The sample includes the Milky Way galaxy and 11 nearby Sb and Sc galaxies with known radial gas surface densities and Hα surface brightnesses, used as a measure of the SFR. The Bayesian approach permits to compute genuine probabilities for the validity of several different functional forms. We do not find any significant explicit dependence of the SFR on radius. A linear dependence on gas density $\Psi \propto g$ is found to be an order of magnitude more probable than any universal non-linear law. If each galaxy has its own exponent x, the non-linear law $\Psi \propto g^x$ would also be highly probable. These findings are fairly independent of the CO-H_2 conversion recipe used.

Key words: Methods: data analysis – Methods: statistical – Stars: formation – Galaxy: structure – Galaxies: ISM – Galaxies: spiral

1. Introduction

The star formation rate (SFR, usually designated as Ψ) specifies how much mass of interstellar gas is converted into stars per unit time. Which physical properties of the interstellar medium it depends on, and what its functional form is – if there is a deterministic relation at all – is still rather unclear. As young objects are found in regions of enhanced gas density, it seems most likely that the SFR increases with the local gas density. From a comparison of the scale heights of young stars and H I gas in the Milky Way, Schmidt (1959, 1963) finds a quadratic dependence on the mass density: $\Psi(z) \propto \rho(z)^2$. From UV photometry Donas et al. (1987) find a linear dependence of the total SFR of a galaxy and the mass of its atomic gas.

Send offprint requests to: J. Köppen[1]

But it is not yet clear whether atomic hydrogen is a good star formation indicator at all: The lack of correlation between total SFR and total H I mass of galactic disks is noted by Kennicutt & Kent (1983) and Balkowski et al. (1986). The inclusion of the molecular component yields an exponent of the density of 1.3 . . . 2.5 (Guibert et al. 1978). Rana & Wilkinson (1986) find a better correlation of the SFR with surface density of molecular hydrogen alone than with either H I or total gas (H I + H_2); the exponent is 1.2 ± 0.2. Studies of the relation of indicators for formation of massive stars (Hα emission, H II regions, blue stars) with gas surface density give a variety of values for the exponent, ranging from 1 to 4 (e.g. Madore et al. 1974, Buat 1992).

There may well be other quantities the SFR depends on: From studies of the Milky Way galaxy and M 83, Talbot (1980) proposes that the SFR in the outer regions of a galaxy is proportional to the frequency with which the gas passes through spiral arms. Wyse & Silk (1989) generalize this approach by also allowing a dependence on the gas surface density. Dopita (1985), Dopita & Ryder (1994), and Ryder & Dopita (1994) suggest a dependence on the *total* mass surface density.

Moreover, there are both observational and theoretical indications for a minimum density or pressure for the SFR (cf. Kennicutt 1989, van der Hulst et al. 1993, Kennicutt et al. 1994, Wang & Silk 1994, Chamcham & Hendry 1996). Kennicutt (1989) finds a SFR which depends nearly linearly on total gas density (with an exponent of 1.3 ± 0.3) as long as the gas surface density exceeds some critical density which is of the order of the critical density for gravitational stability in the disk. A pressure threshold is predicted by Elmegreen & Parravano (1994), so that star formation should cease at nearly three radial scale lengths, whereas the H I disks usually extend beyond this.

The judgment of which law of star formation is the best one depends on how good an agreement can be achieved between observational data and theoretical predictions. Of course, one might get a perfect fit for any data, if only the model had a sufficiently large number of free parameters. Thus, extending a theory by including additional dependencies – which lead to more free parameters – would in general improve the fit, but makes the theory more complex. As the accuracy of the data is finite, one has to make a decision of what level of accuracy of the fit or what level of complexity of the theory is still reasonable. Usually, this decision is made more or less subjectively, based on one's feeling and experience. The metaphor of Occam's razor – of leaving out any unnecessary complexity in a model or theory – is well known as a philosophical concept, but can one devise a practical method that implements this concept when comparing several forms of the law of star formation with the observational data?

In this paper the Bayesian approach to statistics is shown to permit the construction of such a practical method. This method is described, tested with artificial data, and applied to the question of whether the presently available observational data permits a judgement on the form of the SFR. In particular we wish to decide whether the SFR is a power law of the local gas density alone or whether an additional dependence on galactocentric radius exists – or rather to what extend the data permit to make such a decision.

2. Bayesian statistics

In the Bayesian approach to statistics probability is interpreted as a measure of credibility rather than the frequency of occurrence as in the classical approach (see e.g. O'Hagan 1994). This permits to assign a probability for the validity of a physical theory, in our case of a certain law of star formation.

2.1. Probability of a theory

Let us consider a set of N hypotheses H_i, only one of which can be true. Then the probability of the k-th hypothesis, given the data D, is computed from Bayes' Theorem:

$$P(H_k|D) = \frac{P(H_k)\Lambda(D|H_k)}{\sum_{i=1}^{N} P(H_i)\Lambda(D|H_i)} \,. \tag{1}$$

The prior probabilities $P(H_k)$ represent the investigator's degree of belief or his knowledge from previous measurements. The likelihood $\Lambda(D|H_k)$ is a measure of how well the predictions of H_k match the data. Thus, Eq. (1) describes how a measurement or observation improves our knowledge: It states that the *posterior* probability for a certain hypothesis to be true is in proportion to the product of its probability assigned before seeing the data D and the likelihood $\Lambda(D|H_k)$.

In case a hypothesis H_k contains free parameters $\lambda_1, \lambda_2, \ldots$, the above likelihood $\Lambda(D|H_k)$ is replaced by the *mean* likelihood with respect to some prior information about the parameters:

$$\langle\Lambda(D|H_k)\rangle = \int P(\lambda_1,\ldots)\,\Lambda(D|H_k,\lambda_1,\ldots)\,d\lambda_1\,\ldots\,, \tag{2}$$

where $P(\lambda_1, \lambda_2, \ldots)$ is the (joint) prior probability density. Since this prior is normalised over the parameter space,

$$\int P(\lambda_1, \lambda_2, \ldots)\,d\lambda_1\,d\lambda_2 \ldots = 1\,, \tag{3}$$

the *mean* likelihood diminishes, if the parameter space were blown up to a volume much larger than that one where $\Lambda(D|H_k, \lambda_1, \ldots)$ contributes significantly. Likewise, the inclusion of an additional free parameter leads to a smaller mean likelihood. In either case one might thus achieve a better fit of the data, but if one allows oneself too much freedom, this 'dilution' of the *mean* likelihood may more than compensate for any increase of the likelihood itself due to the improved fit. Since in Bayesian reasoning a hypothesis is to be preferred when its posterior probability exceeds that of any other competing hypothesis, it is this feature of 'dilution' which allows a mathematically consistent formulation of Occam's razor.

When further data are available, it can easily be incorporated by applying Bayes' theorem sequentially, using the old posterior probability distributions just for the new prior.

If one has only weak prior information about a parameter there are rules how to construct the prior distribution (Jeffreys 1983). Such a rule ensures for instance that for estimation of parameters the confidence region drawn from the posterior density distribution reproduces the 'classical' confidence region, i.e. given an experiment could be performed very often, the true value of a parameter is within the interval indeed as often as the chosen confidence level states.

In practice, especially for good data, the actual prior may be quite unimportant: the truth has then a good chance against our prejudice (as long as an inappropriate prior does not prevent it at all). Only if the data were so poor or scant that they do not add to our knowledge, the posterior will reflect just our prejudice.

Often one has 'nuisance parameters', i.e. formal parameters in a hypothesis, whose true values are of no interest or relevance, like offsets and factors of proportionality. Then the likelihood is also integrated over the space of this nuisance parameter.

2.2. Simulations with artificial data

As a demonstration of how in the outlined method Occam's razor is at work, and to show some general features, we apply the method to a simpler problem: we wish to decide whether a measured profile of the optical surface brightness $B(r)$ of a galaxy is better represented by an exponential law

$$B_1(r) = c_1 \ \exp(-r/a)\,, \tag{4}$$

with a scale length a and a factor of proportionality c_1, or rather by a more complicated law

$$B_2(r) = c_2 \ r^b \ \exp(-r/a)\,, \tag{5}$$

that has the exponent b as an additional parameter. As parameter space we allow

$$0.05 \leq a \leq 50$$

$$-2 \leq b < 0, \ \ 0 < b \leq 2\,.$$

The factors of proportionality c_k are considered to be nuisance parameters. The 'observational' data are random realizations of

$$B_{\rm obs}(r) = 5 \ r^{-1} \ \exp(-r/5) \cdot 10^{\epsilon}\,, \tag{6}$$

which is sampled at $n = 10$ points at $2 \leq r \leq 12$. The noise ϵ applied is a Gaussian with zero mean and variance σ^2. Then the likelihood is:

$$\Lambda_k = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[\lg\left(\frac{B_{\rm obs}(r_i)}{B_k(r_i)}\right)\right]^2\right)\,. \tag{7}$$

For each of the two laws (4) and (5) we compute the mean likelihood $\langle\Lambda\rangle$ by integrating Eq. (7) over the parameter space. Following Jeffreys, the parameter priors are assumed to be uniform over b, and uniform in the logarithms of amplitude c_k and scale length a. With $0 < c_k < \infty$, the integration over amplitude can be done analytically. This would seem to imply a divergent normalisation integral, but as the same space is used for both laws, the normalisation integral cancels out, because we consider only the ratio of the mean likelihoods. This has been done 50 times with different realizations for the noise ϵ. For the prior probabilities $P(H_k)$, we assume that we have no prior knowledge, and thus both are considered equal.

In Fig. 1 we compare the values of $\lg\langle\Lambda\rangle$ for the two fitting laws (4) and (5), as obtained from the individual data configurations for several values of noise level σ indicated by different symbols. For very low noise (the crowded cloud of open circles), the probability for the more complex law is much larger than that for the simple law. Hence, the true law underlying the data is unambiguously found. As the noise level increases, the cloud of points shows a larger scatter (triangles), and also the mean probability for the complex law drops. At $\sigma = 0.2$ dex the small dots are almost evenly clustered around the diagonal, i.e. on the average both laws are equally probable. Thus it depends on the particular configuration of the observational data which fitting law happens to be come out as the better one. At $\sigma = 0.3$ dex the cloud of crosses has become elongated along the diagonal, but most points

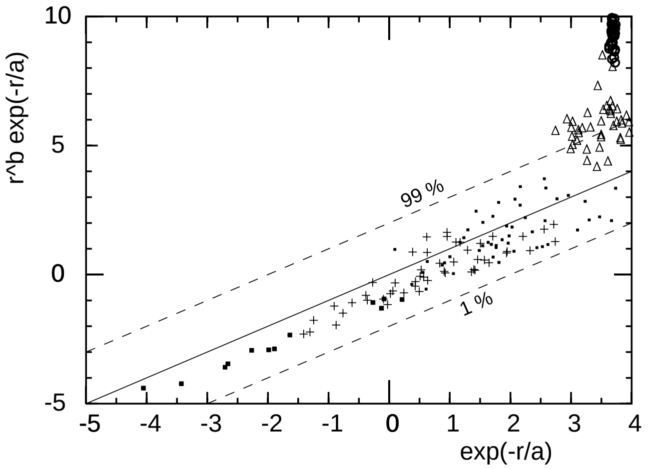

Fig. 1. Simulations with noisy artificial data; shown are the logarithms of the mean likelihoods $\lg\langle\Lambda\rangle$ for a simple law (exponential decay; abscissa) and a more complicated ($r^b\exp(-r/a)$; ordinate) one, for several noise levels: $\sigma = 0.003$ (circles), 0.05 (triangles), 0.2 (small dots), 0.3 (crosses), and 1 dex (black squares). The full line refers to the probabilities being equal, and the dashed lines to the probability levels 1 and 99 per cent

are found in the regime with a higher probability for the simpler law. This is the action of Occam's razor. At still higher levels of the noise, the cloud of black squares is more strongly concentrated towards the diagonal, and in the limit of infinite noise, both laws are equally probable. Then the likelihood is constant in the whole parameter space.

Application of Bayes' Theorem requires that both laws are mutually exclusive ($b \neq 0$), one can assign to both laws probabilities which add up to unity. Points lying above the 99 per cent line mean that the probability of the law (5) is higher than 0.99. That there are no points lying below the 1 per cent line is due to the small parameter space $-2 \leq b \leq 2$. Enlarging this range would make the law (5) more unreasonable, and all points would move down.

This example shows the workings of Occam's razor quite well. It also demonstrates that the scatter in $\langle\Lambda\rangle$ due to different realizations can be quite appreciable. In practice, one often has but one set of data.

2.3. Application to our problem

Now we apply this method to assess models of the SFR. A nested hierarchy of SFR laws is considered, with the most general one being

$$\Psi(g, r) = c\, g^x / r^y\,, \tag{8}$$

with gas surface density $g(r)$ and distance r from the centre. We prefer to explore such an explicit dependence on

r instead of the local angular velocity $\Omega(r) = v_{\rm rot}/r$ or the epicyclic frequency, as this avoids additional errors from the uncertainties of rotation curves. But since many galaxies show a fairly constant rotational velocity, their SFR should well be represented with $y = 1$.

Apart from the most general law (8), several simpler SFR prescriptions $x = 1, 2$ and $y = 0, 1$ as well as combinations are considered, in order to cover already proposed SF hypotheses. Since we shall assume that we have no preference to any one of these laws, we assign equal prior probabilities to all models, irrespective of the number of free parameters.

In what follows the random noise due to observational errors as well as any intrinsic scatter be normally distributed in $\lg\Psi$ with zero mean and variance σ^2. Hence, as suggested by published error bars, it is assumed that the *relative* errors in the SFR as well as in the gas surface densities are normally distributed and do not depend on the brightness level (i.e. constant signal-to-noise ratio). The physical origin of the scatter in the data – observational uncertainties, azimuthal and radial inhomogeneities in the disks, genuine intrinsic spread of the SFR itself – shall not be addressed here, and we shall not make any distinction between the various contributions.

Assuming that the data points are independent from each other, the likelihood is given by:

$$\Lambda(c,x,y,\sigma) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\lg\Psi_{\rm obs}(r_i) - \lg(c\,g(r_i)^x/r_i^y)\right)^2\right), \quad (9)$$

where $\Psi_{\rm obs}$ is the observed SFR indicator, i.e. the Hα surface brightness, and the summation is over all n points. We note that because the errors are assumed to be distributed log-normally, there is no difference between fitting Ψ with g^x and Ψ/g with g^{x-1}. In both ways, we retain the full information.

Unfortunately, information about the error bars of the surface brightnesses in the various wavelengths is not only rather scarce, and is difficult to compare due to the different angular resolutions used, but one must also keep in mind that possibly there is quite a large genuine scatter which is averaged out by the azimuthal integration. In principle a thorough discussion of all the scatter involved should enable one to construct a more informative prior. Since the information necessary is not available, only weak prior information about the true error level is assumed. Thus, we consider σ an additional nuisance parameter, and apply a prior uniform in $\ln\sigma$. The use of a prior uniform in σ does not affect the posterior inference markedly.

The factor of proportionality c, a measure of the efficiency of the star formation, would be a rather interesting parameter to determine, too. This would enable us to compute gas depletion timescales. However, to get this value requires the knowledge not only of the absolute surface brightnesses for all objects, but also corrected for internal extinction in the galactic disk. Since the data comes from different sources and usually not absolutely calibrated, we must consider c here as nuisance parameter, too.

Hence, the likelihood is integrated over c, assuming an uniform prior over $\ln c$, and an interval $c = 0...\infty$. The integration of the likelihood can be performed analytically:

$$\int_0^\infty \Lambda(c,x,y,\sigma)\,\frac{dc}{c} = \frac{\ln 10}{\sqrt{n}\,(2\pi\sigma^2)^{\frac{n-1}{2}}} \exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} Q_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} Q_i\right)^2\right]\right), \quad (10)$$

where $Q_i = \lg\Psi_{\rm obs}(r_i) - x\lg g(r_i) + y\lg r_i$. The normalization integral for this prior diverges for $c \to 0$, but since we consider the ratios of mean likelihoods, and since the same nuisance parameter with the same range is present in all the laws considered, the normalisation factors cancel.

Integrating (10) afterwards with respect to $\ln\sigma$ results in a likelihood $\overline{\Lambda}(x,y)$:

$$\overline{\Lambda}(x,y) = \int_0^\infty \left(\int_0^\infty \Lambda(c,x,y,\sigma)\,\frac{dc}{c}\right)\frac{d\sigma}{\sigma} = \frac{\ln\sqrt{10}}{\sqrt{n\pi^{n-1}}}\,\frac{\Gamma\left(\frac{n-1}{2}\right)}{\left[\sum_{i=1}^{n} Q_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} Q_i\right)^2\right]^{\frac{n-1}{2}}}. \quad (11)$$

Finally, integration over the parameter spaces gives the mean likelihood which is a measure for the posterior probability of a SFR law:

$$\langle\Lambda\rangle = \int\int p_0(x)\,p_0(y)\,\overline{\Lambda}(x,y)\,dx\,dy. \quad (12)$$

$p_0(x)$ and $p_0(y)$ are the *normalised* prior distributions for the model parameters x and y.

Since $\langle\Lambda\rangle$ depends on the volume of the parameter space, the choice of the range for x and y might pose a principal problem: If we really knew nothing about x and y, why not allow an infinitely large space? This could dilute the likelihoods beyond any limit. However, very large exponents (say $|x|, |y| \geq 5\ldots10$) would lead to extremely steep functions and are therefore quite unreasonable. We decided to choose this parameter space:

$$-6 \leq x, y \leq 11,$$

which would contain all reasonable situations, but would not be too restrictive. Within these boundaries the prior density distribution is assumed uniform. If the likelihood 'mountain' occupies only a small area well within this region, any further increase of the parameter space decreases $\langle\Lambda\rangle$ by the factor with which the volume grows.

This analysis is performed for each galaxy separately, yielding for each object $\langle\Lambda\rangle$-values for each of the different SFR prescriptions. The values are suitably normalised to facilitate comparison of the SFR laws, and these Bayes factors are collected in Table 2.

To assess how well the various laws reproduce the data of the complete sample of galaxies, there are two possibilities: The probability for any law, but allowing the best combination of its parameters to vary from object to object, is obtained by multiplying simply the Bayes factors from Table 2. On the other hand, the density distribution for the joint mean likelihood can be computed from all objects, by taking the product over all N galaxies: $\prod_{k=1}^{N} \overline{\Lambda_k}(x, y)$. Integration over the parameter spaces gives the posterior of each law. One obtains the optimal values for the two parameters, and one can extract confidence regions in parameter space.

3. Observational data

From the literature we compiled inclination corrected data on H I mass surface density, velocity integrated CO intensity, and Hα surface brightness, with the latter being considered proportional to the SFR. The Hα line is produced by recombination in H II region ionized by the Lyman-continuum radiation of hot stars, and can thus be considered a most reliable measure of massive star formation (Kennicutt 1983a). If not all stellar photons are absorbed in the H II region gas, this may lead to an underestimation of the SFR (cf. Patel & Wilson 1995a).

The CO brightnesses are converted into H_2 mass surface densities by the same factor for all the galaxies: $3.68\,\mathrm{M_\odot\,pc^{-2}/(K\,km\,s^{-1})}$ in accordance with Gordon & Burton (1976), whose data is used in Smith et al. (1978). Usually CO was measured only in a smaller, interior portion of the disk for which Hα-data is available. Then we consider only that part of the disk for which both data are available. The mass surface density of all the gas is obtained from that of hydrogen by multiplication of a factor 1.4, to account for a helium fraction of ten per cent (by number). All the data is linearly interpolated onto the grid given by the Hα data.

We do not wish to make any further selection of the data which might reflect some underlying prejudice. Hence, we include all points, for which all three data are available. In the following we give the references and notes for individual galaxies. The types are from Sandage & Tammann (1981):

- Milky Way galaxy: All data are from Smith et al. (1978) who give for H I and H_2 the masses in annuli. The central region with $R/R_\odot < 0.8$ is neglected here, because the uncertain amount of non-thermal contribution, as well as the outer region $R/R_\odot > 1.3$, where the data are scarce. We have taken their SFR (derived from the galactic distribution of young radio H II regions as measured at 5 GHz) directly in the place of Hα brightness. The problem of dust obscuration, which is a severe one in Hα studies, does not exist for these radio selected sources.
- NGC 2841 (Sb): H I from Bosma (1981), CO from Young & Scoville (1982), Hα from Kennicutt (1989), Hα starts at 2.7 kpc.
- NGC 4254 = M 99 (Sc(s)I.3): H I and CO from Kenney & Young (1988a), Hα from Kennicutt (1989).
- NGC 4303 = M 61 (Sc(s)I.2): H I data from Warmels (1988), CO surface brightness from Kenney & Young (1988b), Hα from Kennicutt (1989),
- NGC 4321 = M 100 (Sc(s)I): H I from Warmels (1988), CO from Kenney & Young (1988b), Hα from Kennicutt (1989).
- NGC 4535 (SBc(s)I.3): H I from Warmels (1988), CO from Kenney & Young (1988b), Hα from Kennicutt (1989).
- NGC 4654 (SBc(rs)II): H I from Warmels (1988), CO from Kenney & Young (1988b), Hα from Kennicutt (1989).
- NGC 4689 (Sc(s)II.3): H I from Warmels (1988), CO from Kenney & Young (1988b), Hα from Kennicutt (1989).
- NGC 4736 = M 94 (RSab(s)): H I from Garman & Young (1986), CO from Kenney & Young (1988b), Hα from Kennicutt (1989).
- NGC 5194 = M 51 (Sbc(s)I-II): H I and CO from Scoville & Young (1983), Hα from Kennicutt (1989).
- NGC 5457 = M 101 (Sc(s)I): all data from Kenney et al. (1991).
- NGC 6946 (Sc(s)II): H I and CO from Tacconi & Young (1986), Hα from Kennicutt (1989), except for the point with extremely low Hα at 330 arcmin (his Fig. 2) which is ignored and the data linearly interpolated from the adjacent points.

4. Results for individual galaxies

4.1. The confidence regions

Figure 2 shows for all twelve galaxies the confidence regions in the parameter space: Each line depicts the contour level enclosing the area within which 90 per cent of the contributions to $\langle\Lambda\rangle$ is found. Their modes, i.e. the best values for the exponents x and y, are marked in the figure and are given in Table 1.

The plot shows that the parameter modes cluster around a linear or quadratic dependence on gas density, and there is little indication in favour of an explicit dependence on radial distance: The horizontal line $y = 0$ crosses or touches (for NGC 4321) the 90 per cent regions of all galaxies but the Milky Way galaxy.

Except for NGC 2841, the orientation of the confidence ellipses is quite similar. This is merely a consequence of the radial profile of the gas density: Let $\Psi(r) \propto r^a$ and $g(r) \propto r^b$. Since we test $\Psi(g, r) \propto g^x r^{-y}$, a good fit must have: $y = bx - a$. Usually the gas density decreases outward ($b < 0$, see Fig. 7) which gives rise to the anti-correlation seen in Fig. 2.

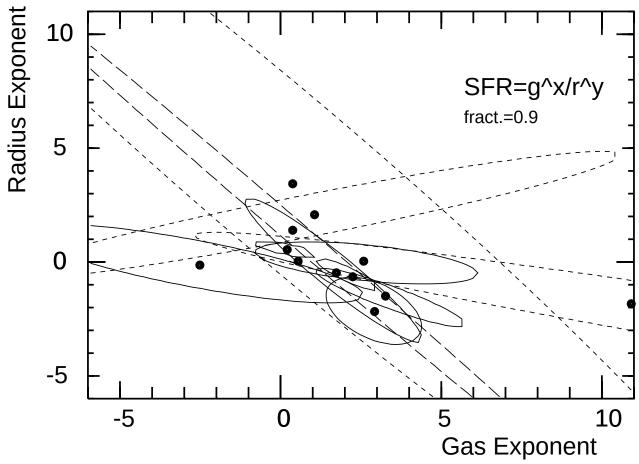

Fig. 2. The locations in the parameter space of the modes (black dots) and the contours within which 90 percent of the total integrated $\langle\Lambda\rangle$ lies. The dashed lines refer to galaxies whose likelihood peak lies outside the region shown (NGC 4689) or whose 90 per cent region is partially cut

Table 1. The best parameters for fitting $\Psi \propto g^x/r^y$

Galaxy	x	y
Milky Way	2.93	−2.18
NGC 2841	1.06	2.08
NGC 4254	2.25	−0.65
NGC 4303	3.27	−1.50
NGC 4321	0.21	0.55
NGC 4535	−2.52	−0.14
NGC 4654	2.59	0.04
NGC 4689	$\geq$ 11.00	$\leq$ −1.84
NGC 4736	0.38	3.44
NGC 5194	1.74	−0.48
NGC 5457	0.54	0.00
NGC 6946	0.38	1.40

4.2. Comparison of SFR laws

We consider eight laws for the SFR. Three have no free parameters, viz. the linear and quadratic dependence on gas density and $\Psi \propto g/r$. Four have one free parameter, such as $\Psi \propto g^x$. Finally there is the most general law with two free parameters: g^x/r^y. Since all the laws are nested into the most general one, we can compute their mean likelihoods by integrating the likelihood mountain over the appropriate parameter space, weighted with the relevant prior distribution for the parameters.

Bayes' theorem (Eq. 1) applies only to mutually exclusive hypotheses. This means that in a more general law, e.g. g^x, we must exclude all sub-hypotheses it contains, here: g and g^2. If one considers a sub-hypothesis of a lower dimension as a separate hypothesis, one takes that into account in the parameter prior distribution by superposing a δ-distribution peaked at each parameter value characterising a sub-hypothesis. Thus, exclusion of the sub-hypotheses is simply done by integration with the unmodified parameter prior (cf. O'Hagan 1994).

Afterwards, the posterior probability of each law is computed by multiplying with the law's prior probability. Since we do not give any preference to any one of the laws, we assign the same prior value. We emphasise that in doing so, we actually treat the 'simpler' laws less favourably than we would be tempted from everyday practice, where one would give preference to a law with fewer free parameters. In Table 2 we compare, for the sake of convenience, only appropriately normalized Bayes factors, not the real $\langle\Lambda\rangle$s, for each galaxy.

In Sect. 2.2 $\langle\Lambda\rangle$ has been shown to depend strongly on the outcome of a realization in the presence of appreciable noise. In order to check the robustness of our results we generated for each galaxy in the sample new data sets by leaving out every one datum at a time, i.e. from n data we make n sets of $n-1$ data points. Computing the values of $\langle\Lambda\rangle$ for the two SFR-laws $\Psi \propto g^x$ and $\propto g^x/r^y$ gives an impression of the sensitivity due to the choice of data points. Table 3 provides the means and the dispersions of $\lg\langle\Lambda\rangle$ and of the best parameters for both star formation laws. In Fig. 3 the two values of $\lg\langle\Lambda\rangle$ obtained from each configuration are depicted.

In half of the galaxies the scatter is lower than about 0.35 dex, thus we expect the real uncertainty for a given Bayes factor to be about a factor of two or three. The worst cases are the Milky Way galaxy and NGC 4321. For the Galaxy, the large scatter is due to the innermost and outermost points. Neglecting either one would improve the goodness of fit drastically. In NGC 4321 the large dispersion is due to the innermost point. We notice that removal of the innermost or the outermost datum point would strongly increase $\langle\Lambda\rangle$ in eight of the twelve galaxies of our sample. This may reflect merely the fact that often the profile of the Hα surface brightness shows a steep drop towards the centre, and in some galaxies near the outer rim, too. For his analysis Kennicutt (1989) excluded the innermost 2 kpc, arguing that there the extinction in H II regions can be considerably higher than in the disk proper, or even the star formation law itself could be disturbed from its normal form. While this exclusion of rim points may strongly improve $\langle\Lambda\rangle$, tests show that the overall assessment of the SFR laws is not strongly influenced. For instance, the optimal parameter values for the joint mean likelihood – see Fig. 4 – are still within the 90 per cent confidence region, if the rim points are left out.

4.3. Comments on individual galaxies

For several galaxies, some remarks are necessary:

Table 2. Individual Bayes factors for 8 models of the star formation rate SFR, depending on gas surface density g and galactocentric radius r. For each galaxy, the factors are normalized to the most probable model

SFR:	g	g^2	g/r	g^x	g^x/r	g/r^y	g^2/r^y	g^x/r^y
Milky Way	0.182	0.328	0.024	0.042	0.004	0.079	1	0.289
NGC 2841	0.019	0.001	0.392	0.235	1	0.453	0.435	0.262
NGC 4254	1	0.189	0.001	0.187	0.001	0.037	0.661	0.043
NGC 4303	1	0.168	0.013	0.096	0.012	0.094	0.423	0.226
NGC 4321	0.499	1	0.007	0.098	0.267	0.786	0.025	0.299
NGC 4535	0.108	0.056	0.023	1	0.111	0.045	0.029	0.076
NGC 4654	0.206	1	0.020	0.172	0.006	0.012	0.034	0.008
NGC 4689	0.621	1	0.336	0.406	0.165	0.063	0.078	0.159
NGC 4736	0.569	1	0.952	0.362	0.378	0.472	0.460	0.322
NGC 5194	1	0.268	0.355	0.089	0.052	0.089	0.102	0.018
NGC 5457	0.001	0.000	0.000	1	0.000	0.006	0.000	0.008
NGC 6946	0.371	0.417	1	0.080	0.082	0.099	0.094	0.057

Table 3. Results from multiple analyses of the data, leaving out one datum at a time: the averages and dispersions for lg $\langle\Lambda\rangle$ and for the parameters

Galaxy	SFR $\propto g^x$		SFR $\propto g^x/r^y$		
	lg $\langle\Lambda\rangle$	x	lg $\langle\Lambda\rangle$	x	y
Milky Way	-3.65 ± 0.60	1.9 ± 0.2	-2.79 ± 1.01	2.7 ± 0.6	-1.9 ± 0.9
NGC 2841	-0.52 ± 0.26	-5.5 ± 0.3	-0.47 ± 0.22	0.7 ± 2.5	2.0 ± 0.8
NGC 4254	-1.29 ± 0.39	1.4 ± 0.1	-1.90 ± 0.33	2.1 ± 0.4	-0.5 ± 0.3
NGC 4303	-0.38 ± 0.42	1.6 ± 0.2	-0.65 ± 0.58	2.4 ± 1.4	-0.5 ± 1.5
NGC 4321	-2.36 ± 0.62	1.6 ± 0.2	-2.16 ± 0.33	0.3 ± 0.2	0.5 ± 0.1
NGC 4535	0.70 ± 0.29	-3.0 ± 0.7	0.00 ± 0.15	-3.3 ± 2.0	0.1 ± 0.8
NGC 4654	0.37 ± 0.32	2.6 ± 0.4	-0.72 ± 0.28	2.2 ± 1.0	0.2 ± 0.5
NGC 4689	-1.47 ± 0.40	3.5 ± 1.5	-1.95 ± 0.38	≥ 11	≤ -1.8
NGC 4736	-1.91 ± 0.34	3.3 ± 0.5	-1.93 ± 0.32	0.2 ± 4.5	4.2 ± 5.6
NGC 5194	-2.97 ± 0.26	1.3 ± 0.1	-3.58 ± 0.27	1.7 ± 0.5	-0.5 ± 0.8
NGC 5457	5.27 ± 0.25	0.54 ± 0.00	3.35 ± 0.42	0.54 ± 0.07	0.02 ± 0.05
NGC 6946	-1.00 ± 0.48	1.5 ± 0.2	-1.14 ± 0.45	0.4 ± 1.1	1.3 ± 1.3

- The Milky Way galaxy: Using all points, one gets a nearly circular confidence ellipse, as shown (cf. Fig. 2), and for the most probable parameter set ($x = 2.93$, $y = -2.18$) a rather poor fit with a strong radial trend of the residuals. The results depend strongly whether one excludes the innermost and/or outermost points at $R/R_\odot = 0.24$ and 1.25. This is evident already from Fig. 3, where eight of the ten data points cluster at the lower left edge of the diagram. The two remaining points, with much higher probabilities, result from taking away either the inner or the outer point. Excluding both, a much better fit could be obtained, without any radial trend of the residuals. The confidence region would become more elongated, like for the other objects, and the linear Schmidt law would now be well included. This law would be 25 times more probable than the one-parameter law g^x.
- NGC 4254: The radial profiles of H I, CO, and Hα are very smoothly dropping with increasing radius. Thus a narrow confidence region is produced. Only Hα shows an increase in the inner 2 kpc. If one left out this central region, the goodness of fit for a non-linear Schmidt law could be improved considerably, with no radial dependence being necessary.
- NGC 4321 shows a central enhancement of the Hα brightness. Without this central point the linear Schmidt law would be roughly 25 times more probable than the non-linear Schmidt law.
- NGCs 4535 and 4654 have fairly large ellipses, mainly because of the small number of points (4 and 5).
- NGC 4689: From the seven data points one obtains a very strongly elongated confidence ellipse, the maximum probability density is outside the parameter space. The innermost point has a lower Hα brightness than the next one further out. If one removed this point, the ellipse would remain as elongated, but the maximum density would happen to be at the opposite boundary of the parameter region.

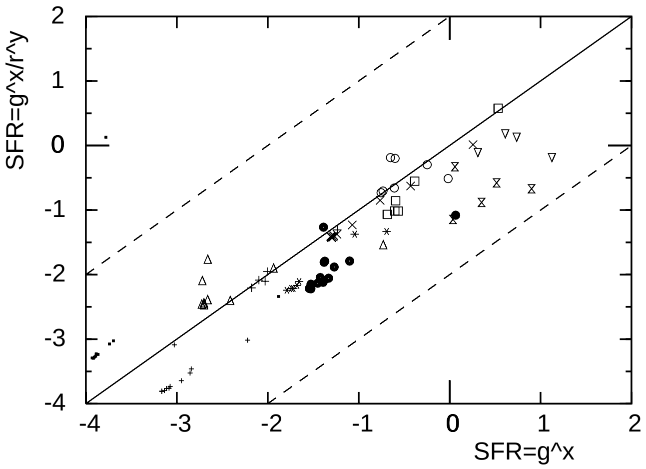

Fig. 3. The probabilities $\lg \langle \Lambda \rangle$ for two hypotheses for the SFR law, as computed for each galaxy by using all but one datum. Along the solid line the probabilities for the laws are equal, along the dashed lines they are 1 and 99 per cent or vice versa. The symbols designate: Milky Way (small filled square), 2841 (open circle), 4254 (filled circle), 4303 (open square), 4321 (delta), 4535 (nabla), 4654 (hourglass), 4689 (asterisk), 4736 (+), 5194 (smaller +), 6946 (x). For convenience, NGC 5457 is not shown, as its points cluster at much higher values (5,3)

- NGC 4736 has the largest confidence region of the sample, thus nearly all SFR laws are equally probable. Its Hα brightness profile is characterized by an inner ring about 100 times brighter than the exterior part. To explain such a strong variation with the rather smooth H I and CO-profiles gives rise to a rather poor fit for any combination of the parameters.
- NGC 5194: eleven points yield a fairly large ellipse. While both H I and CO show a smooth radial profile, Hα exhibits a rather strong bump between 5 and 7 kpc, which is difficult to explain by a simple, systematic dependence on gas density or radius.
- NGC 5457: eleven data points with a very smooth Hα-profile yield by far the smallest confidence region, dominating thereby the joint probability distribution.
- NGC 6946: The eleven data points give a strong correlation between x and y, because accidentally the run of total gas surface density with radius follows a power law quite closely (cf. Fig. 7). The ratio of Hα and CO brightnesses is almost constant with radius, Hα showing more structure than the very smooth CO profile.

5. Joint probabilities

As already discussed in Sect. 2.3 one can consider two cases for the joint probability for a law to be valid in all the sample objects: The free parameters

- are allowed to vary from galaxy to galaxy. This is done by taking the *product* of the individual Bayes factors from Table 2 for all objects.
- vary for all galaxies simultaneously. This is obtained by taking the product of the likelihood densities $\overline{\Lambda_k}(x, y)$ to get the density distribution for the *joint* mean *likelihood.*

The resulting values are presented in Table 4, normalised to the largest value. By far the most probable laws are the power law dependence on gas density – with the exponent being different for each galaxy – and the linear law $\Psi \propto g$. The other parameter-free laws can definitely be excluded. The one-parameter laws have factors which are at least one order of magnitude lower than those for the linear law, with g^x and g/r^y being by far more probable than the others. Of the remaining laws, g^x/r can safely be excluded. Interestingly, the law g/r^y does rather well; probably this is because deviations from the average, if concentrated near the centre, can be covered by this law very well by taking small positive or negative y. The most general law, with two parameters, is unlikely, but not the worst one.

One notices that the products usually give lower probabilities than the ones derived from the joint likelihood. This is because the two approaches have different sensitivities to a change of the size of the parameter space: In the first approach, the normalisation of the parameter priors enters in each individual Bayes factor. If e.g. for all twelve objects the main contributions to the mean likelihood were contained in the parameter space, doubling of its volume would result in a reduction of the product by a factor 2^{12}. The probabilities from the joint likelihood would be merely halved.

The confidence regions obtained from the joint likelihood are presented in Fig. 4, using a smaller parameter space in order to bring out the details. The standard factor for the conversion of measured CO surface brightness into H_2 surface density (cf. next Section) is assumed. Notice that such a region must not be an ellipse and may even consist of two disconnected regions. The tongue at the lower left is due to the increase of the density towards $x = 0.5$ and $y = 0$ caused by NGC 5457.

One obtains a region which is rather narrowly ($\Delta x \approx 1$ and $\Delta y < 1$) confined around a nearly linear dependence of the SFR on gas density, and independent of radial distance. The best value for the gas exponent and the confidence interval for $y = 0$ agree well with the findings of Kennicutt (1989). The linear law, $\Psi \propto g$, is well within this 90 per cent confidence region, and thus has one of the largest joint Bayes factors (cf. Table 4). From both approaches for the joint probability, we conclude that with the standard conversion factor *there is no need to consider any explicit radial dependence of the SFR.*

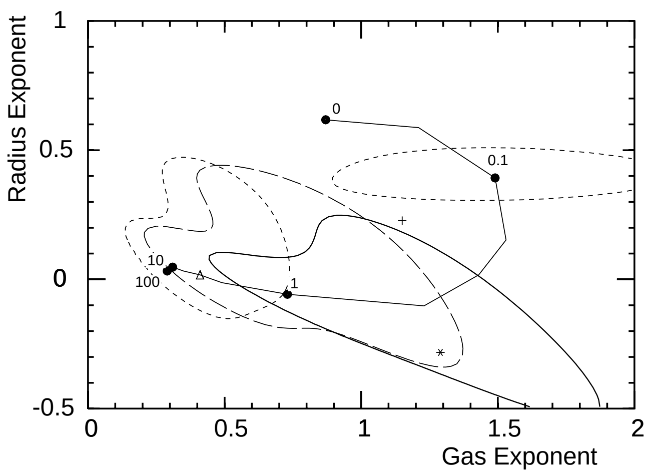

Fig. 4. The 90 per cent confidence region of the joint likelihood for all galaxies and how it depends on the CO-H_2 conversion factor X. The solid line depicts the region computed with the standard value $X = 3.68$, short dashes indicated the regions for $X/3.68 = 0.1$ and 100, and long dashes refer to application of the metallicity-dependent prescription of Arimoto et al. (1996). Filled circles show the best parameters obtained for different values of $X/3.68$. For the plus-sign the Galaxy and NGC 5457 are excluded, and the asterisk indicates the parameters for all galaxies if data of the rim points are removed (cf. Sect. 4.2)

Table 4. Joint Bayes factors, from the product of individual Bayes factors and from the joint likelihood density

SFR-law	product	joint likelihood
g	1	
g^2	$1.9 \cdot 10^{-6}$	
g/r	$1.3 \cdot 10^{-15}$	
g^x	0.59	0.061
g^x/r	$2.0 \cdot 10^{-13}$	$4.1 \cdot 10^{-12}$
g/r^y	$1.5 \cdot 10^{-5}$	0.033
g^2/r^y	$4.5 \cdot 10^{-7}$	0.002
g^x/r^y	$9.8\, 10^{-6}$	0.002
$x_{\rm best}$	—	0.73
$y_{\rm best}$	—	−0.06

6. The CO-H_2 conversion factor

There is agreement that CO is a good tracer for molecular hydrogen, but the value of the conversion factor (from velocity integrated CO intensity into molecular hydrogen surface mass density) – and whether a common factor is appropriate at all – is still rather unclear (cf. Bloemen 1989, Vila-Costas & Edmunds 1992, Boselli et al. 1995, Sodroski et al. 1995, Arimoto et al. 1996). The value $X = 3.68$ from Gordon & Burton (1976) still serves as a widely used 'standard' value which seems to be appropriate at least for giant molecular clouds (Magnani & Onello 1995).

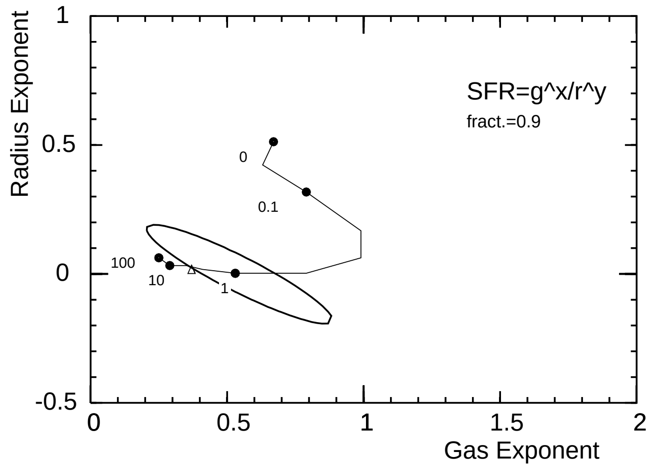

Fig. 5. Positions of the most likely parameters of the law $\Psi \propto g^x/r^y$ in NGC 5457, as the CO-H_2 conversion factor X is changed. The numbers denote $X/3.68$, its ratio relative to the standard value; the kinks depict the ratios 0.03, 0.2, 0.3, 0.5, 2, and 5. The ellipse is the 90 per cent confidence ellipse as in Fig. 2. The triangle shows the position obtained by using the conversion recipe of Arimoto et al. (1996)

6.1. Changing X for the whole sample

We investigate the effect of changing this factor more closely, because the molecular gas is thought to be more intimately related to star formation than the atomic one (cf. Sect. 1). Hence, raising X a better fit might be expected. Fig. 4 shows the path of the peak of the likelihood mountain in the x-y-plane when changing the conversion factor from 0 (i.e. the gas responsible for star formation is only atomic hydrogen) to very large values (i.e. only molecular hydrogen). An increase by a factor 2 or a decrease by a factor 3 takes the maximum of the joint likelihood outside the confidence region for the standard factor. If one considered only molecular gas $X \gg 1$, the maximum becomes sharper, but still is rather close to the standard confidence region. On the other hand, if only atomic hydrogen were considered, i.e. $X = 0$, one needs an explicit radial dependence of the SFR. This is because in most galaxies the radial H I profile is flatter than that of the CO-surface brightness. This also shows that star formation seems more closely linked to the molecular gas than to the atomic gas.

Contrary to the expectation, we find formally a maximum of $\langle\Lambda\rangle$ between one fourth and one third of the standard value of the conversion factor, although this is not a very pronounced one. $\langle\Lambda\rangle$ is increased by about a factor of 4 which is not significant in view of the scatter seen in Table 3.

The Milky Way galaxy and NGC 5457 (M 101) are somewhat different from the other galaxies, as they are the only ones where the H I, CO, and Hα brightnesses are given for the same radial rings. Furthermore, obtaining the radially dependent data for the Galaxy requires additional assumptions than the more directly available data from external galaxies. Finally, NGC 5457 has a confidence ellipse very much smaller than that of any other object. The dramatically larger probability density makes this galaxy dominate in any kind of joint probability. But even if one excludes both objects from the sample, the confidence region remains of about the same size but is shifted somewhat towards the upper left, the maximum lying at the point marked with a plus-sign. This is still within the confidence region for the sample as a whole. There is a very strong overlap, thus the results are not critically dependent on the inclusion of NGC 5457 and its narrow peak.

6.2. NGC 5457

This galaxy has by far the smallest confidence ellipse in the x-y parameter space (cf. Fig. 2). This is partly due to the fact that data is given at 20 radial points, and partly because all the profiles look rather smooth. Thus, this galaxy is particularly well suited to study the influence of the conversion factor. Figure 5 shows how the peak of the joint likelihood moves across the x-y-plane as X is changed. It is worth emphasizing that, just like for the joint likelihoods of the entire sample, changing $X/3.68$ between 0.5 and 10 one still remains within the 90 per cent confidence region. With the metallicity-dependent conversion factor the maximum is just outside the region, at $x = 0.37$ and $y = 0.02$, i.e. with almost no radial dependence. $\langle\Lambda\rangle$ is largest for $X \to \infty$, about a factor 3 larger than for the standard value $X = 3.68$. This means that the best fit is achieved for relating the SFR with the molecular gas rather than the atomic gas.

Instead of considering a variation of X one can introduce a weight factor w which describes the relative importance of the two gas species for the star formation process. Hence, the total gas density is substituted by the weighted mean:

$$g(r) = (1-w)\,\Sigma_{\rm HI} + w\,\Sigma_{\rm H_2}\,. \qquad (13)$$

Performing the integration over w (with a constant prior) results in the likelihood distribution shown in Fig. 6. The peak is situated at $x = 0.3$ and $y = 0.05$, but a linear Schmidt law with no dependence on radius is still within the 90 per cent confidence contour. The linear Schmidt law is found to have the highest probability, though it is slightly outside the 90 per cent r egion for the parameter of the non-linear one: $0.25 \le x \le 0.75$. This illustrates nicely Occam's razor: the parameter space for the non-linear Schmidt law is just too large.

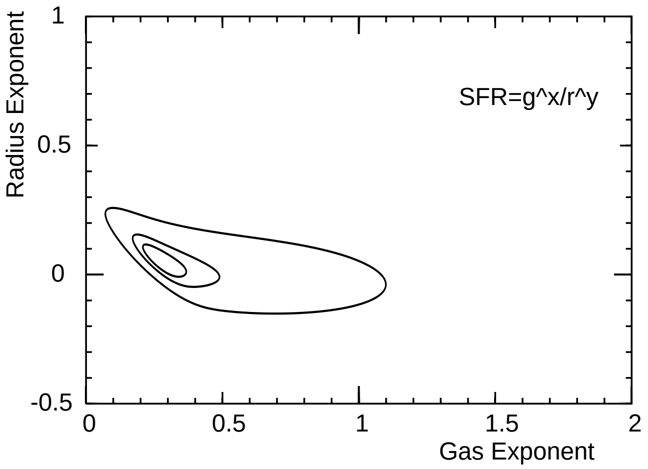

Fig. 6. The confidence contours for 90, 50, and 25 per cent for the law $\Psi \propto g^x/r^y$ in NGC 5457, with the CO-H_2 conversion factor X being a nuisance parameter

6.3. A metallicity dependent conversion factor

From CO-observations of individual molecular clouds, Arimoto et al. (1996) derive a conversion factor that depends on the local oxygen abundance of the gas. Such a dependency would allow to take into account the claimed underestimate of the relative amount of molecular gas at the outskirts of the Milky Way galaxy (Lequeux et al. 1993, Sodroski et al. 1995) as well as in low-luminosity galaxies (Boselli et al. 1995). For 7 galaxies of our sample, including the Milky Way galaxy, radial abundance gradients are known. Taking the corresponding abundance data from Vila-Costas & Edmunds (1992), we re-analyze the objects of this subsample. The results are collected in Tables 5 and 6. Only for the Milky Way galaxy, the value of $\langle\Lambda\rangle$ is significantly (factor 10) increased over that derived with the standard conversion factor. Likewise, Fig. 4 shows that the confidence region for the joint likelihood strongly overlaps with the one from the standard assumption. Thus, it seems that in general there is no urgent necessity to apply a more sophisticated conversion recipe: Table 6 does not significantly deviate from Table 4.

7. Conclusions

The Bayesian approach to statistics allows to compute genuine probabilities for the credibility of a law, incorporating not only the goodness of its fitting of given observational data but also any a priori knowledge or preference on its validity. This permits to develop a statistically consistent formulation of what is known as Occam's razor,

Table 5. Same as Table 1, but using the metallicity-dependent CO-H_2 prescription of Arimoto et al. (1996). The last column shows the relative increase in $\langle\Lambda\rangle$ as compared with the standard case $X = 3.68$

Galaxy	x	y	$\langle\Lambda\rangle / \langle\Lambda\rangle_{(X=3.68)}$
Milky Way	2.24	−3.02	10.6
NGC 4254	1.40	−0.30	0.6
NGC 4303	2.59	−1.50	1.2
NGC 4321	0.20	0.54	0.9
NGC 5194	1.56	−0.31	1.0
NGC 5457	0.38	0.04	2.3
NGC 6946	1.56	−0.30	0.9

Table 6. As Table 4, but using the metallicity-dependent CO-conversion recipe

SFR-law	product	joint likelihood
g	0.24	
g^2	$1.4 \cdot 10^{-12}$	
g/r	$3.0 \cdot 10^{-17}$	
g^x	1	0.010
g^x/r	$2.6 \cdot 10^{-14}$	$7.9 \cdot 10^{-12}$
g/r^y	$1.1 \cdot 10^{-5}$	0.005
g^2/r^y	$9.4 \cdot 10^{-8}$	$1.2 \cdot 10^{-7}$
g^x/r^y	$3.0 \cdot 10^{-4}$	$2.9 \cdot 10^{-4}$
$x_{\rm best}$	—	0.38
$y_{\rm best}$	—	0.04

that a simpler law should be preferred over an unnecessarily complicated theory.

We develop such a practical approach and demonstrate its working by simulations with artificial data. We apply this method to the question, which kind of dependence of the star formation rate on gas surface density and galactocentric radius represents best the presently available observational data of galactic disks. As indicator for the SFR the Hα surface brightness is used, and the gas density g is obtained from both atomic and molecular gas. The SFR laws of the type $\Psi \propto g^x/r^y$ are considered, both in its general form with two free parameters and restricted forms with one or no free parameter. This is done for twelve galaxies, including the Milky Way. For the individual objects as well as for the joint probabilities of the entire sample, one gets the largest probabilities for the linear 'Schmidt law' $\Psi \propto g$ and for the more general law g^x where the exponent x may be different for different galaxies. The probability for other forms including the quadratic Schmidt law g^2, is much lower. In particular, the accuracy of the data presently rule out th need for any explicit dependence on the galactocentric radius, This finding agrees with the global studies of Donas et al. (1987) and Buat (1992) as well with the local one by Kennicutt (1989).

As the test example shows, one has to keep in mind that the given observational data are just one realization of the data, which may give rise to an appreciable amount of scatter in the Bayes factors. In our study, we do not separate the various sources of observational uncertainties and genuine scatter in the physical processes. The dependence of the results on which radial points are included, poses some problem. For several galaxies one finds that excluding a datum at the inner or outer edge may greatly improve the fit and thus causes the scatter of the value of $\langle\Lambda\rangle$ in Table 3 (cf. Sect. 4.3). However, in the overall assessment, e.g. by the joint likelihood, inclusion of these 'disturbing' points does not alter the basic finding that it is the simple linear Schmidt law which is most compatible with all the data. Finally, this finding is also quite insensitive to changes of the CO-H_2 conversion factor, and even the employment of a metallicity-dependent conversion prescription does not change the situation.

This might mean that the observational data is still so poor or so strongly dominated by genuine irregularities that all what one can say, is that the SFR increases about linearly with the gas density, i.e. that the inverse SF timescale Ψ/g is constant. Because of the assumption that the *relative* errors follow a normal distribution, our method would not detect an intrinsic radial dependence of Ψ/g, if it were compensated by a power-law dependence of the gas surface density g on radius. Among the galaxies in the sample, such a situation arises only once, with NGC 6946. The radial profiles of the (atomic and molecular) gas are shown in Fig. 7, where only NGC 6949 exhibits the straight line characteristic for a power-law. But since most galaxies follow rather closely an exponential decrease of $g(r)$ – examples are NGCs 4254, 5194, 5457 – such a compensation is not *common* to all objects. This is a strong argument in favour that SFR and SF timescale do not depend on radial distance explicitly.

While this suggests that the projected SFR Ψ increases linearly with the gas surface density, does this really exclude a non-linear relationship between the SFR *per volume* ψ and the *volume* density ρ of interstellar matter? A counterexample may illustrate that this must not be the case: Consider a vertical column with cross-section A in the galactic disk. Suppose that all the gas is in the form of N clouds with a certain internal density $\rho_{\rm c}$ but different volumes: $V_1, V_2, \ldots V_N$. Star formation also occurs only in these clouds, with the SFR per volume $\psi = k\rho^n$. Summing over all clouds gives the SFR per unit area $\Psi = k \sum_{i=1}^N \rho_{\rm c}^n \cdot V_i/A$, and the surface density $\Sigma = \sum_{i=1}^N \rho_{\rm c} \cdot V_i/A$. Therefore, the SFR per unit area may be proportional to the gas surface density

$$\Psi = k\,\Sigma\,\rho_{\rm c}^{n-1}\,,$$

even if $n \neq 1$, only provided that the internal density of the star forming clouds $\rho_{\rm c}$ is constant. It goes without saying that a study of the relation between averaged quantities can never substitute spatially resolved studies.

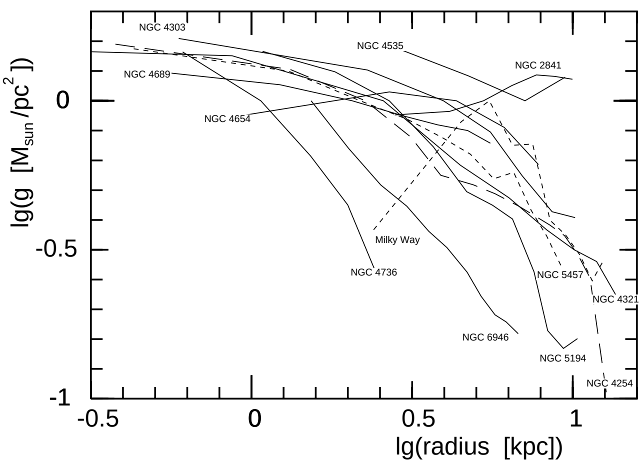

Fig. 7. The radial profiles of the gas surface density – with the standard CO-H_2 conversion factor. For clarity, some curves are shown dashed.

Acknowledgements. J.K. thanks the Astrophysikalisches Institut Potsdam for invitations to Babelsberg during which most of this work was done. We also thank the Deutsche Forschungsgemeinschaft for financial support (project He 1487/13-1).

References

Arimoto N., Sofue Y., Tsujimoto T., 1996, PASJ **48**, 275
Balkowski C., Alloin D., LeDenmat G., 1986, A&A **167**, 223
Bloemen H., 1989, ARAA **27**, 469
Boselli A., Gavazzi G., Lequeux J., et al., 1995, A&A **300**, L13
Bosma A., 1981, AJ **86**, 1791
Buat V., 1992, A&A **264**, 444
Casoli F., Dickey J., Kazès I., et al., 1996, A&A **309**, 43
Chamcham K., Hendry M.A., 1996, MNRAS **279**, 1083
Dopita M.A., 1985, ApJ **295**, L5
Dopita M.A., Ryder S.D., 1994, ApJ **430**, 163
Donas J., Deharveng J.M., Laget M., Milliard B., Huguenin D., 1987, A&A **180**, 12
Elmegreen B.G., Parravano A., 1994, ApJ **435**, L121
Garman L.E., Young J.S., 1986, A&A **154**, 8
Gordon M.A., Burton W.B., 1976, ApJ **208**, 346
Guibert J., Lequeux J., Viallefond F., 1978, A&A **61**, 1
Jeffreys H., 1983, Theory of Probability, 3rd ed., Clarendon Press, Oxford
O'Hagan A., 1994, Kendall's Advanced Theory of Statistics, Vol. 2B, Bayesian Inference, Edward Arnold, London
Kenney J.D., Young J.S., 1988a, ApJ **326**, 588
Kenney J.D., Young J.S., 1988b, ApJ Suppl.Ser. **66**, 261
Kenney J.D.P., Scoville N.Z., Wilson C.D., 1991, ApJ **366**, 432
Kennicutt R.C., 1983a, ApJ **272**, 54
Kennicutt R.C., 1983b, A&A **120**, 219
Kennicutt R.C., 1989, ApJ **344**, 685
Kennicutt R.C., Kent S.M., 1983, AJ **88**, 1094
Kennicutt R.C., Tamblyn P., Congdon C.W., 1994, ApJ **435**, 22
Lequeux J., Allen R.J., Guilloteau S., 1993, A&A **280**, L23
Madore B.F., van den Bergh S., Rogstad D., 1974, ApJ **191**, 317
Magnani L., Onello J.S., 1995, ApJ **443**, 169
Patel K., Wilson C.D., 1995a, ApJ **451**, 607
Patel K., Wilson C.D., 1995b, ApJ **453**, 162
Rana N.C., Wilkinson D.A., 1986, MNRAS **218**, 497
Ryder S.D., Dopita M.A., 1994, ApJ **430**, 142
Sandage A., Tammann G.A., 1981, A revised Shapley-Ames catalog of bright galaxies, Washington
Schmidt M., 1959, ApJ **129**, 243
Schmidt M., 1963, ApJ **137**, 758
Scoville N.Z., Young J.S., 1983, ApJ **265**, 148
Smith L.F., Biermann P.B., Mezger P.G., 1978, A&A **66**, 65
Sodroski T.J., Odegard N., Dwek E., et al., 1995, ApJ **452**, 262
Tacconi L.J., Young J.S., 1986, ApJ **308**, 600
Talbot R.J., 1980, ApJ **235**, 821
van der Hulst J.M., Skillman E.D., Smith T.R., et al., 1993, AJ **106**, 548
Vila-Costas M.B., Edmunds M.G., 1992, MNRAS **259**, 121
Wang B., Silk J., 1994, ApJ **427**, 759
Warmels R.H., 1988, A&AS **72**, 427
Wyse R.F.G., Silk J., 1989, ApJ **339**, 700
Young J.S., Scoville N.Z., 1982, ApJ **260**, L41

This article was processed by the author using Springer-Verlag LaTeX A&A style file *L-AA* version 3.