

## PART I

### Empirical Studies



## 2. Technological paradigms and the evolution of networks: lessons from the pharmaceutical industry

**Fabio Pammolli and Massimo Riccaboni**

---

### INTRODUCTION

This chapter studies the evolution of a set of interorganizational relationships in a domain characterized by structural breakthroughs in underlying knowledge bases: the international pharmaceutical industry in the last 20 years.

Sectoral specificity notwithstanding, our work unravels some general and distinctive properties of the relationships between patterns of technological change and the evolution of organization forms, particularly in terms of the analytical strategies they call for.

Our starting points are the notion of ‘technological paradigms’ (Dosi, 1982) and a focus on markets and organizations as repositories of problem-solving knowledge (see Simon, 1962, 1991). Moreover, we refer to the notion of ‘appreciative theorizing’, according to which theorizing tends to be close to empirical work and provides both guidance to and interpretation of empirical investigations (see Nelson and Winter, 1982).

In accordance with the interplay between theory and data that should sustain any appreciative theorizing exercise, our empirical work is exploratory in nature. We use graph-theoretical tools and measures to reveal how the nature and evolution of relevant technological conditions induce distinguishable patterns of structural change in the set of economic relationships we analyse (see Orsenigo et al., 2000; Riccaboni, 2000; Pammolli et al., 2001). Since relationship-induced effects are not negligible relative to effects associated with other variables, we take into account the overall relational structure. Specifically, we associate the network with a digraph (directed or oriented graph), by identifying the actors with the vertexes of the graph and the relationships with the edges.

It is our claim that the topological methods of graph theory deal appropriately with the evolving nature of networks. First, facing a system of

contractual relationships that is perturbed by technological breakthroughs, they encode all relevant information on the global structure of the set of relationships. On the contrary, any approach focused on variables defined at the level of individual nodes cannot convey information on the global structure of the system, since it is focused on local properties and local evolution. Second, a graph-theoretical analysis of the system at different points in time (comparative static) can enlighten the transition between different relational regimes. Namely, qualitative changes in the structure of the network are represented by topologically non-equivalent graphs. Finally, a graph-theoretical approach – as every topological method – extracts relevant information on the evolution of the system. As a consequence, graph-theoretical tools are both conceptually simpler and computationally cheaper than methods based on differential equations in a finite-dimensional space.

The chapter is organized into three sections. In the first section we analyse the evolution of the network, investigating an extensive data set that covers more than 5000 research licensing agreements of known technological content between around 2000 firms for 1978–1997. We refer to the notion of ‘Canonical decomposition’ of a graph in order to disentangle two major drivers or components of the structural evolution of the network. That is, co-specialized and transversal actors that rely on co-specialized and general purpose research technologies. In the second section, the structural information extracted from the topological analysis is used to build a non-parametric statistical test on the nature of the connection between technological bases and relational behaviours within the system. The final section sums up the main findings and implications of our analysis.

## THE STRUCTURE OF THE NETWORK

The last 25 years have witnessed a revolution in biological sciences, with significant basic advances in molecular biology, cell biology, biochemistry, protein and peptide chemistry, physiology, pharmacology and other relevant scientific disciplines. The application of these new bodies of knowledge to the pharmaceutical industry has had an enormous impact on the nature of R&D activities, on the organizational capabilities required to introduce new drugs, and on patterns of industry evolution. In particular, it has been emphasized that the emergence of a dense set of collaborative relationships among firms of different types and other research institutions has been a major feature of the recent evolution of the industry.

In synthesis (a far more detailed account can be found in Orsenigo et al.,

2000 and in Riccaboni, 2000), the recent evolution of research strategies and technologies in pharmaceutical R&D can be characterized by discerning between two main regimes that coexist and complement each other within the industry. The first regime is based on research techniques and hypotheses that stay coupled with specific biological hypotheses and are specific to given fields of application (co-specialized technologies), while the second regime is characterized by new generic research tools and techniques (general purpose technologies). In the case of co-specialized research technologies, the design and experimentation of each new drug requires individual analysis, and lessons learned from the design and experimentation of one therapeutic cannot immediately be transferred to the development of other classes of drugs. Conversely, the appearance of general purpose technologies for the production and screening of new molecular structures has introduced a new dimension in the organization of the relevant knowledge bases, since the new technologies tend to be applicable to multiple biological targets and diseases.

In the empirical analysis that follows, an association is made between research hypotheses and techniques, and R&D projects. Moreover, every firm is defined by the collection of its research projects over time, while licensing agreements are conceived as contractual and organizational devices through which hypotheses and techniques are combined. We focus on agreements drawn before the start of clinical development activities. Finally, since we deal with licensing agreements, an ‘Originator’ can always be distinguished from a ‘Developer’.

This section analyses in detail the transformations that occurred in the organization of innovative activities within the international pharmaceutical industry between 1978 and 1997.

Data used for this study are drawn from the Pharmaceutical Industry Database (PHID) at the University of Siena. PHID integrates information from several sector-specific fonts including a proprietary database on more than 14000 R&D projects and, for collaborative agreements, Bioscan, Recombinant Capital, IBI and Pharmaventures databases, annual reports (SEC files) and specialized press news (Scrip, Spectrum). PHID provides information on the typology, technological content and date of signing for 2600 R&D licensing agreements signed by 1461 firms and institutions<sup>1</sup> during the pre-clinical development phase. Specifically, our sample includes 349 leading pharmaceutical firms and 1112 new biotechnology firms.

In the context of this essay, these data are represented by means of a directed graph.

**Definition 1:** A directed graph (digraph, oriented graph) is the datum  $G = (V, E, f, s)$  where:

- $V$  and  $E$  are non-empty sets (whose elements are called respectively vertexes and edges of the graph)
- $f$  and  $s$  are functions from  $E$  to  $V$  such that:
  - $f(e) \neq s(e)$  for every  $e \in E$  and
  - $f(e_i) = f(e_j)$  and  $s(e_i) = s(e_j) \Rightarrow e_i = e_j$  for every  $e_i, e_j \in E$

Roughly speaking, we can conceive a digraph as a set of points (vertexes) linked by edges; for each  $e \in E$  the vertexes  $f(e)$  and  $s(e)$  can be thought respectively as the starting and the ending point of the edge. In this representation, conditions (1) and (2) forbid self-loops and parallel edges respectively.

More precisely, the following identifications are established:

1.  $V$ : Firms and institutions with at least one R&D project in their pipelines.
2.  $E$ : Pharmaceutical R&D projects included in the data set.
3. Orientation: for every R&D project, we are able to single out the Originator (that is, the company or institution that started the R&D project) and the Developer. We then identify the starting and ending point of an oriented edge respectively with the Originator and the Developer of the corresponding project.

As we are dealing with the structural evolution of the network, we focus on projects for which an Originator can be identified and distinguished from a Developer. This explains why in the corresponding digraph self-loops are forbidden.

According to the above definitions, the structure of the network can be represented, at any given time  $\tau$ , by means of a directed graph  $M_\tau(E, V)$ , where  $V$  is the set of vertices and every edge  $e$  within the graph is an oriented link defined by a couple  $(o, d)$ , (Diersel, 1997). Alternatively, the directed graph  $M_\tau$  can be represented by an ‘adjacency matrix’  $M_\tau \Leftrightarrow A(M_\tau) = [a_{do}]$ . Matrix entry  $a_{do}$  equals 1 if an edge  $e(d, o)$  does exist and 0 otherwise.

Moreover, we introduce a function  $t: V \rightarrow \mathfrak{R}$ , called the entry time function, defined in the following way:

$$t(v) = \text{time of } v \in V \text{ entry into the market}$$

This function induces a ‘total ordering’ on the set of vertexes, which we refer to as the ‘chronological labelling’ of vertexes.<sup>2</sup>

Back to the matrix notation, we can permute the adjacency matrix in order to obtain the ordered matrix:  $A(M_\tau) = [a_{do}]_{\leq \tau}$ , where  $\{t(1) \leq \dots \leq t(d)$

$\leq \dots \leq t(n) \leq \tau\}$ , and  $\{t(1) \leq \dots \leq t(o) \leq \dots \leq t(m) \leq \tau\}$ , with  $m$  and  $n$  representing the total number of Originators and Developers active within the market at time  $\tau$ . Sometimes – for the sake of synthesis – we shall refer to the block matrix  $B(G)_{\leq \tau(\theta)}$  obtained after collapsing rows and columns of matrix  $A(M_{\leq \tau})$  that correspond to firms or institutions belonging to a common cohort of entrants  $\theta = [\tau, \tau + \delta)$  (Generation) for a suitable  $\delta > 0$ . In that case, entries  $b_{do}$  of  $B(G)_{\leq \tau(\theta)}$  indicate the total number of contractual relationships between Generations  $d$  and  $o$ .

In sum, the network is referred to as a directed graph (digraph), as for every R&D project it is possible to distinguish the Originator from the Developer. In addition, the set of vertexes of the digraph has been ordered according to dates of foundation and/or dates of entry within the market.

As a result, two distinct time-dimensions have been identified: the first is defined at the project level (the Originator–Developer distinction) and corresponds to the orientation of the graph; the second is singled out at the macro level, and it is related to the time-evolution of the graph.

The importance of the technological determinants of the structural evolution of the network can be appreciated, at a first glance, by looking at the cumulative number of agreements, classified according to firms' dates of entry. Our analysis (see Orsenigo et al., 2000) reveals that:

1. Originators have entered the network by introducing successive waves of new research technologies, which shape the overall evolution of the network;
2. firms already active within the network have not played a major role as Originators in the new technological trajectories that emerged after their entry;
3. earlier entrants have gained access to the new technological trajectories mainly as Developers;
4. as time goes by, the rate of entry in any given technological trajectory has slowed down. That is to say, entrants are closely linked to the generation of new technological trajectories.

Evidence on patterns of entry, relational roles of earlier and later entrants and, finally, on new technological waves, suggest the existence of a dynamic process with the following properties. Major new technological breakthroughs initially induce the entry of new firms, which act as specialized technology Originators. As time goes by, Developers succeed in developing internal capabilities in the new fields. Correspondingly, relational intensity, as well as flows of entry, shift forward to new technologies and firms;

5. After 1992, the emergence of 'transversal technologies' like combinatorial chemistry has changed the structure of the network. New

entrants based on new general purpose technologies have started to act as Originators of projects licensed to different types of firms, irrespective of age.

In this chapter, we want to characterize how different combinations of actors and relational roles within the network affect its global structure at different points in time. In order to do that, a Canonical Dulmage–Mendelsohn decomposition has been performed (see Dulmage and Mendelsohn, 1958, 1959).

We have applied a condensation procedure to the bipartite graph, generating a graph minor  $bG_{\Delta_t}[M]$ , which has been obtained by shrinking every strongly connected subgraph, replacing it with a vertex, and then substituting each set of parallel lines with single lines.

In the case of a bipartite graph, the concept of a strongly connected component is equivalent to that of a strong Hall component – a component that has a perfect matching, that is, a matching, which covers every vertex within it (for further details, see Diersel, 1997). The lines that belong to a matching are said to be admissible, while the remaining ones are called inadmissible.<sup>3</sup>

The application of a Canonical Dulmage–Mendelsohn decomposition to the condensed bipartite graph  $bG_{\Delta_t}[M]$  (see Lovasz and Plummer, 1986, p. 137) produces the following results:

1. Two subgraphs  $H_1, H_2$ , which are the connected components of the induced subgraph  $bG_{\Delta_t}[M]$ .
2.  $H_1, H_2$  are two elementary bipartite graphs.
3. Since the number of connected components of  $bG_{\Delta_t}[M]$  is greater than one, by permuting rows and columns the corresponding bi-adjacency matrix  $A(bG_{\Delta_t})$  can be put into the form:

$$\begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix}$$

where matrices  $A_1, A_2$  are the square bi-adjacency matrices corresponding to the subgraphs  $H_1, H_2$ , while \* represents ties that are transversal with respect to the two submatrices.

In the Canonical Dulmage–Mendelsohn decomposition, a major role is assigned to the interplay between maximum matching and minimum vertex covers. As transversal vertices are included in every minimum vertex cover of the graph, a greater proportion of such kind of vertices out of the total number of Originators would imply a higher number of time chains in which the graph can be decomposed.

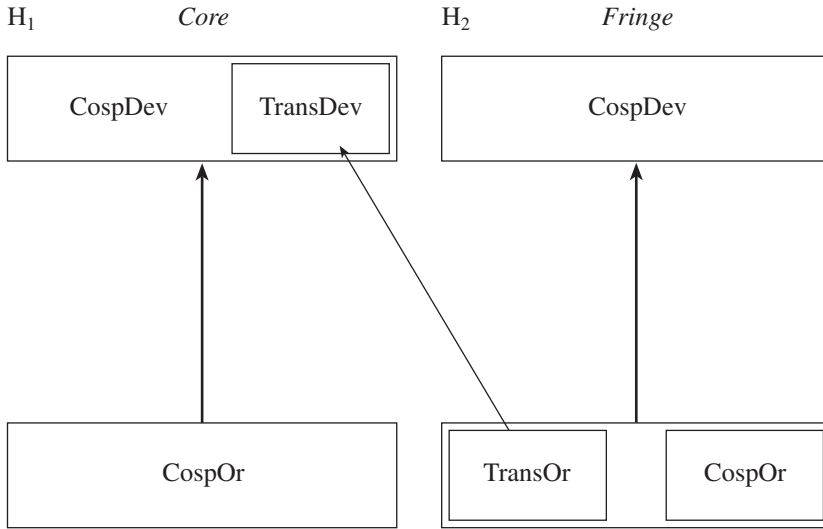


Figure 2.1 Transversal and co-specialized nodes within the graph

Figure 2.1 synthesizes the logic and the result of the Dulmage–Mendelsohn decomposition. Boxes  $H_1$ ,  $H_2$  represent the two non-trivial subgraphs for which a matching can be found. In each box, we observe two subsets of Developers and Originators. Box  $H_1$  contains the relational core of the market (approximately, the persistent relational component of the graph: that is, firms which have a large number of agreements and/or have entered the market early on), while box  $H_2$  includes the relational fringe of the graph. The matching in box  $H_1$  captures the main structuring process. However, we also identify a subset of Developers in box  $H_1$  that link with a subset of Originators in box  $H_2$ . They correspond to what we defined above as ‘Transversal Developers’ (*TransDev*) and ‘Transversal Originators’ (*TransOr*).

The two sets of firms denoted as *TransDev* and *TransOr* can be thought of as the structural attractors of the market, that is, they attract most of the agreements in each period of time (they are present in all the intersections among minimum coverage vertex sets).

*TransDev* and *TransOr* firms cannot be assigned an unambiguous relational role within the market, that is, they play a transversal role. The core group of Developers (*TransDev*) establishes several relationships with a wide variety of firms. On the other side, within the Originators group, a clear distinction can be drawn between a set of firms that are co-specialized in their relational behaviour (*CospOr*), that is, they are matched, and a set of firms that play a transversal role within the market (*TransOr*).

The comparison of the output of the procedure at different points in time shows unequivocally that a variety of generative processes and corresponding relationships has characterized the evolution of the graph. In particular, starting from the beginning of the 1990s, a new technological paradigm driven by the entry of new general purpose research technologies has started to couple with the previous one, mainly based on the growth through specification of biological knowledge. The coupling between the two regimes has induced a dramatic increase of the overall degree of interdependence within the network. That is to say, the network has started to be more and more indecomposable, because of the coexistence of a variety of decomposition criteria, integrative mechanisms and, moreover, because of a higher degree of interdependence among research trajectories. Further information on the technological bases of relational transversality has been gained through a detailed analysis of the technological background of Transversal Originators based on personal interviews, information provided by 10K and 10Q SEC files reports, specialized press, and our proprietary data set on R&D projects within the industry. As a result, we are able to state that relational roles that have been identified within the network correspond to firms embodying different types of technologies. In particular, our controls show that firms specialized in transversal research technologies tend to play a transversal role within the network. Almost all the firms which are included in PHID as active in general purpose research technologies (that is, new drug delivery systems, combinatorial chemistry, genomics, genomic libraries, proteomics, high-throughput screening and bioinformatics) have been classified by the algorithm as Transversal Originators. Conversely, and most important, all the firms that have been identified as Transversal Originators into the graph by means of our analytical procedures embody Transversal Technologies. As an example of the implications of the emergence of the new transversal component within the network, Figure 2.2 shows the results of the canonical decomposition of the graph referred to the set of agreements signed in 1997, after the emergence of the transversal component of the network. Submatrices  $A_1$ ,  $A_2$  correspond to subgraphs, while  $T$  marks the matrix area in which transversal R&D collaborations are confined. The area  $T$  contains nodes (firms) which are transversal within the graph.

## TECHNOLOGICAL BASES AND RELATIONAL BEHAVIOUR

In this section we detect how structural variables affect the probability of a tie being present or absent within the graph, by means of a logit

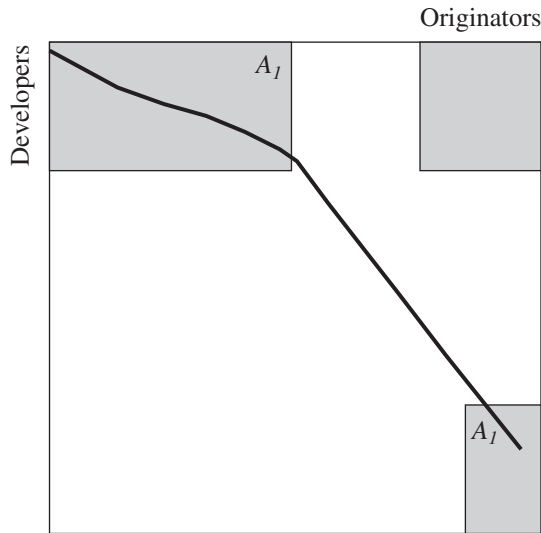


Figure 2.2 Canonical decomposition of the graph, 1997

Pseudolikelihood (p-star,  $p^*$ ) model. Logit  $p^*$  models are non-parametric models that analyse the presence or the absence of a tie going from  $j$  to  $i$  as the dichotomous stochastic variable  $x_{ij}$  conditionally dependent on the presence or absence of other ties in a given 'neighbourhood'. Since every tie is both dependent on its neighbourhood ties and part of the neighbourhood of other ties, the graph instantiates a local self-organizing process given by the joint probability of particular local substructures occurring.

It is worth noticing that the following analysis would have been meaningless in the absence of our previous explorative inspection of the structural properties of the graph. First, although economic theory can offer guidance in the identification of the most meaningful parametric model among the multitude of possible ones, this condition rarely holds in fields which are new and characterized by processes of radical technological change. In such circumstances, purely econometric approaches can be both not identified and self-confirming. As a consequence, a different strategy was followed in this chapter, with an inductive structural analysis coming first and aiming at unravelling major empirical regularities and conditioning variables. Moreover, we take into account the potential effects of an inadequate control upon population heterogeneity on structural estimates. In particular a wide-ranging analysis of the determinants of structural variation and heterogeneity has been undertaken in advance, to boost and explain the fitting outputs of our statistical tests.

The first  $p^*$  model was described in Strauss and Ikeda (1990), and the specification of the relevant neighbourhood was based on Markov graphs, which were meant to be generalizations of other models of interactions first developed in spatial studies (see Kindermann and Snell, 1980), and in physics (see Besag, 1974). As a matter of fact, the Markovian assumption greatly simplifies the dependence structure within the graph and leads naturally to logit parametric models. In the context of this work, it is important to remember that logit  $p^*$  models rest upon two crucial assumptions:

- Assumption 1: Markov property: Edges in the graph are interdependent if and only if they have an actor in common. In the context of our analysis, the Markov Property implies that the probability of a given firm to subscribe two different licenses, given the overall set of licenses, is not simply equal to the product of their marginal conditional probabilities. The Bernoulli model, in which all edges are conditionally independent, can be considered as an extreme case, and more complex kinds of local structural interdependence can be taken into account.
- Assumption 2: Homogeneity property: All the ties within the graph have the same dependence structure. This second assumption rules out heterogeneity in the dependence structure, since it implies that the probability of observing a tie depends on the same local graph configuration irrespectively of differences among agents.

The plausibility of assumptions 1 and 2 has to be evaluated in the context of the present work. First of all, it is reasonable to assume that a firm's licensing activity depends on its strategy and technological background, and not on other firms' behaviour. The licenses of a given firm are interdependent, as they draw on the same technological basis and are coherent with the same general strategy. As a consequence, a Markovian assumption is more plausible than the independence one, while it does not imply a strong restriction as compared to a hypothesis of general interdependence. On the contrary, the homogeneity assumption appears to be particularly strong and restrictive, suggesting an extremely cautious and careful interpretation of the regression outcomes.

In fact, in order to obtain informative results and appropriate conditioning information, we must take into account the effect of heterogeneous technological capabilities embedded into firms and influencing market structure and dynamics.

With both the Markov and the homogeneity assumptions holding at the highest level of aggregation, the general logit  $p$ -star model can be written as:

$$\varpi_{ij} = \log \left\{ \frac{\Pr(X_{ij} = 1 | X_{ij}^c)}{\Pr(X_{ij} = 0 | X_{ij}^c)} \right\} = \theta' [z(x_{ij}^+) - z(x_{ij}^-)]$$

where  $z(x_{ij}^+)$  is the value of a given vector of graph statistics with  $x_{ij}$  set equal to 1 and  $z(x_{ij}^-)$  is the realization of  $z$  with  $x_{ij}$  set to 0. We calculated the Maximum Pseudo-Likelihood Estimations of parameters  $\theta$  for different graph change statistics  $d(x_{ij}) = z(x_{ij}^+) - z(x_{ij}^-)$ .

The importance of our previous structural analysis is clear from the analysis of the output of two different relational models (Tables 2.1 and 2.2). The models are based on three parameters: *Cos* (co-specialization),

Table 2.1 1st model (without generation blocks)

Par.	b	S.E.	Wald	exp(b)
<i>Cos.</i>	-7.1519	0.0400	31919.99	0.0008
<i>Gen.</i>	0.0781	0.0013	3712.16	1.0812
<i>Syn.</i>	0.1840	0.0051	1288.07	1.2021

Notes: -2 log likelihood: 17618.80  
 Goodness of fit: 531635.25  
 Model chisquare: 964046.73 df: 3  
 Overall fit: 99.80  
 Residual (absolute): 2758.80  
 Residual (squared): 1403.91

Table 2.2 2nd model (with generation blocks)

Par.	Block	b	S.E.	Wald	exp(b)
<i>Cos.</i>	-	-6.8537	0.0364	35506.39	0.0011
<i>Gen.</i>	<i>Pharma</i>	0.0615	0.0034	319.57	1.0635
<i>Gen.</i>	NBF<81	0.0808	0.0041	396.91	1.0841
<i>Gen.</i>	NBF82-89	0.0804	0.0023	1180.02	1.0837
<i>Gen.</i>	NBF90-97	0.0807	0.0019	1758.94	1.0840
<i>Syn.</i>	<i>Pharma</i>	0.3024	0.0120	451.37	1.2673
<i>Syn.</i>	NBF<81	0.2250	0.0113	398.32	1.2524
<i>Syn.</i>	NBF 82-89	0.2074	0.0106	386.21	1.2304
<i>Syn.</i>	NBF90-97	0.1378	0.0101	187.97	1.1478

Notes: -2 log likelihood: 17897.27  
 Goodness of fit: 622573.47  
 Model chisquare: 963768.26 df: 8  
 Overall fit: 99.80  
 Residual (absolute): 2747.59  
 Residual (squared): 1397.97

*Gen* (generality), and *Syn* (synthesis). The parameters refer to three change statistics which are respectively edges, 2-out stars and 2-in stars. These measures have been selected among the stars of size three or less defining the neighbourhood of each node, as they are the most significant in terms of both Wald and likelihood ratio tests. The co-specialization parameter measures the effect of one-to-one relationships within the net and substantiates the tendency toward co-specialized research activity, while the other two parameters denote the transversality effect both for Originators (*Gen*) and Developers (*Syn*), measured through 2-out and in stars.

Tables 2.1 and 2.2 summarize the results of our econometric estimates. Notably, the large, negative and significant values of *Cos* in both models constitute reliable measures of the importance that coupled firms had within the net over time. Even more interestingly, the other two parameters are both positive, signifying that transversality, in particular as Developers (*Syn*), leads to higher probability to draw licenses. All in all, these results are coherent with the previous analysis in this field but they are far more accurate in terms of structural insight. Moreover, while the model presented in Table 2.1 does not take into account any heterogeneity among actors, Table 2.2 shows how the three parameters vary among different groups of firms. In this second model, a block structure is defined on the basis of available information on sources of heterogeneity among firms. In particular, we distinguish pharmaceutical firms and NBFs founded before 1981, from 1982 to 1989, and after 1990. As a result of the introduction of a generational structure in our analysis, the values of *Syn* become higher for pharmaceutical firms and older NBFs, than for younger NBFs. On the other hand, parameter *Gen* is almost the same for all NBFs, but now the standard error is lower for younger generation of NBFs. The lower value of *Gen* standard error after 1990 means that a high fraction of younger NBFs have a generalist attitude, as compared to the very few of previous generations.

The econometric exercise that we have accomplished shows the importance of structural information generated by topological methods to sustain and complement any representation of relational intensity based on differential operators. At the same time, our econometric analysis conveys some additional information, which could not be extracted otherwise. All in all, a comparison between the two specifications of our model reveals that, in the absence of the structural analysis and the identification of key heterogeneities among firms, any econometric estimate would hide a fundamental distinction between old and new generalists. On this, it is important to notice that from our qualitative substantive knowledge of the industry we know that while a few of the first generation originators came to be transversal thanks to the cumulateness of the molecular biology

regime, transversality of younger originators is a consequence of the paradigm shift associated with the introduction of the new general purpose research techniques (see Pammolli et al., 2001).

## CONCLUDING DISCUSSION

Even in the neo-Schumpeterian and evolutionary tradition, where an overwhelming amount of research has been devoted to the study of the effects of technological change on the organization of industries, the conceptual problems associated with the nature of technological change and its influence on organization forms have tended to induce a dichotomy between qualitative case studies and conventional analytical methods or modelling strategies.

While from a conceptual point of view it is clear that an evolutionary approach should deal with interdependencies among varieties of decompositions and relevant subsystems, the empirical analysis of the interdependencies that arise within and among organizations in response to the features of technological evolution is far from methodological accuracy and rigour.

In particular, an important issue which arises thus far refers to the problem of admissible aggregations and degrees of abstraction in the study of dynamic interdependent systems (Simon and Ando, 1961; Ando and Fisher, 1963). In the class of systems characterized by Simon and Ando (1961), the interactions between groups were non-null and arbitrary, but weak if compared with interactions within groups. Simon and Ando labelled this class of systems as 'Nearly Completely Decomposable Systems' (NCDS), proving that, for NCDS, the analysis of the short-run and the long-run dynamics of the system can be separated without loss of generality. Afterwards, Courtois (1977) developed both a measure of the degree of approximation and a sufficient condition for admissible aggregation, based on the degree of coupling between subsystems and on the indecomposability of each subsystem.

The Courtois sufficient condition for near-complete decomposability can be stated as follows:

$$\gamma < \frac{[1 - \max_c |\lambda^*(2_c)|]}{2}$$

where  $\gamma$  is the maximum value of the probability to observe off-diagonal entries, and  $(1 - |\lambda^*(2_c)|)$  is the lower bound of aggregate block diagonal matrix  $C$  indecomposability, defined in function of the largest second eigenvalue of the diagonal matrix,  $\lambda^*(2_c)$ .

As a first approximation, the Courtois condition can be restated in terms of degrees of transversality and matching, or coupling, within the graph. However, this measure refers to a matrix of transition probabilities among different states of a given system, which has a given and stable structure, whereas in our empirical analysis we were facing a matrix representing the state of system interdependencies which was changing its own structure in time. As a consequence, the state transition matrix of the graph must be analysed. In particular, our comparative static exercises based on the Canonical Dulmage–Mendelsohn decomposition have shown that when the level of  $\gamma$  rises up, the system does not change its decomposability properties in a continuous fashion. For this reason, it was important to have a topological inspection preceding and complementing any analytical exercise.

In conclusion, we want to state that the graphical toolkit we have introduced is useful to capture the essence of dynamics when systems are far from stable and unambiguous equilibrium forms. We expect that such an apparatus, if further developed and refined, can deserve insightful applications in other domains, whenever structural breakthroughs, multiple orderings and radical technological change are important issues.

## NOTES

1. Every agreement may include different contract typologies. Information on the technological content is available for every agreement, as it refers to the underlying discovery technology. Mergers and acquisitions have been taken into account by collapsing the information relative to the firms engaged in consolidation deals.
2. This total ordering is different from the partial ordering corresponding to the orientation of the graph. On the relationships between order set theory and graph theory see Asratian et al. (1998), Chapter 10.
3. An edge  $e$  is inadmissible if and only if there exists a ‘minimum vertex cover’— that is, a cover consisting of as few elements as possible —  $C_o \subset V_o$  of vertices in  $V_d$  (and vice versa), such that  $e$  belongs to that cover ( $e \in E(G[C])$ ) (see Lovasz and Plummer, 1986; Asratian et al., 1998).

## REFERENCES

- Ando, A. and Fisher, F.M. (1963), ‘Near-decomposability, partition and aggregation, and the relevance of stability discussions’, *International Economics Review*, 4: 53–67.
- Asratian, A.S, Denley, T.M.J. and Haggkvist, R. (1998), *Bipartite Graphs and their Applications*, Cambridge: Cambridge University Press.
- Besag, J.E. (1974), ‘Spatial interaction and the statistical analysis of lattice systems (with discussion)’, *Journal of the Royal Statistical Society, Series B*, 34: 75–83.

- Courtois, P.J. (1977), *Decomposability: Queuing and Computer System Applications*, New York: Academic Press.
- Diersel, R. (1997), *Graph Theory*, Berlin: Springer-Verlag.
- Dosi G. (1982), 'Technological paradigms and technological trajectories. A suggested interpretation of the determinants and directions of technical change', *Research Policy*, 11: 147–62.
- Dulmage, A.L. and Mendelsohn, N.S. (1958), 'Some generalization of the problem of distinct representatives', *Canadian Journal of Mathematics*, 10: 230–41.
- Dulmage, A.L. and Mendelsohn, N.S. (1959), 'A structure theory of bipartite graphs of finite exterior dimensions', *Transactions of the Royal Society Canada Ser.*, 53: 1–13.
- Frank, O. and Harary, F. (1982), 'Cluster inference by using transitivity indices in empirical graphs', *Journal of the American Statistical Association*, 77 (380): 835–40.
- Frank, O. and Strauss, D. (1986), 'Markov graphs', *Journal of the American Statistical Association*, 81: 832–42.
- Harary, F., Norman, R.Z. and Cartwright, D. (1975), *Structural Models: An Introduction to the Theory of Directed Graphs*, New York: John Wiley & Sons.
- Kindermann, R.P. and Snell, J.L. (1980), 'On the relation between Markov random models and social networks', *Journal of Mathematical Sociology*, 8: 1–13.
- Lovasz, L. and Plummer, M.D. (1986), 'Matching theory', *Annals of Discrete Mathematics*, 29, Amsterdam: North Holland.
- Nelson, R.R. and Winter S.G. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, MA: Harvard University Press.
- Orsenigo, L., Pammolli, F. and Riccaboni, M. (2000), 'Technological change and network dynamics. Lessons from the pharmaceutical industry', *Research Policy*, 30: 485–508.
- Pammolli, F., Riccaboni, M. and Orsenigo, L. (2001), 'Variety and irreversibility in scientific and technological systems: the evolution of an industry network', in Pagano, U. and Nicita, A. (eds), *The Evolution of Economic Diversity*, London: Routledge, pp. 216–39.
- Riccaboni, M. (2000), 'Network evolution in drug discovery', PhD dissertation, Sant'Anna School of Advanced Studies, Pisa.
- Simon, H.A. (1962), 'The architecture of complexity', *Proceedings of the American Philosophical Society*, 106: 467–82.
- Simon, H.A. (1991), 'Organizations and markets', *Journal of Economic Perspectives*, 5: 25–44.
- Simon, H.A. and Ando, A. (1961), 'Aggregation of variables in dynamic system', *Econometrica*, 29: 111–38.
- Strauss, D. and Ikeda, T. (1990), 'Pseudolikelihood estimation for social networks', *Journal of the American Statistical Association*, 85: 204–12.